European School of Molecular Medicine

Università degli studi di Napoli Federico II

PhD in Molecular Medicine XXVI Cycle

Molecular Oncology

A Novel Long Non-Coding RNA Transcript Regulates The Fate of Colon Cancer Initiation

Tutor:
Prof. Francesco Salvatore

PhD candidate:
Serena Di Cecilia

Internal Supervisor:
Prof. Luigi Del Vecchio

External Supervisor:
Prof. Martin J. Walsh

Coordinator: Prof. Francesco Salvatore

Academic Year: 2013-2014

To Francesco, Mayra & Imma:

may the curiosity be your lifeblood,
because what you learn today,
for no reason at all,
will help you discover
all the wonderful secrets
of tomorrow.

TABLE OF CONTENT

FIGURE INDEX

TABLE INDEX

LIST OF ABBREVIATIONS

APC	Adenomatous Polyposis Coli
CCICs	Colorectal Cancer Initiating Cells
CCND1	Cyclin D1
CCSCs	Colorectal Cancer Stem Cells
CD	cluster of differentiation
CICs	Cancer Initiating Cells
CK	Cytokeratin
CLIP	Cross-Linking and Immuno-Precitation
CRC	Colon and Rectal Cancer
CSCs	Cancer Stem Cells
ENCODE	Encyclopedia of DNA Elements
FACS	Fluorescence-activated Cell Sorting
FISH	Fluorescence *In Situ* Hybridization
FSC	Forward Scatter
GO	Gene Ontology
lincRNAs	large intergenic non-coding RNAs
LINE	Long Interspersed Elements
lncRNA	long non-coding RNA
LRP6	Low density lipoprotein Receptor-related Protein 6
LUST	LUca-15 Specific Transcript
miRNAs	microRNAs
MYC	Myelocytomatosis oncogene cellular homolog
ncRNAs	Non-Coding RNAs
NMD	Nonsense-Mediated Decay
qRT-PCR	quantitative Reverse-Transcriptase Polymerase Chain Reaction
RBM5	RNA-Binding Motive 5
RIP	RNA Immuno-Precipitation
SINE	Short Interspersed elements
SSC	Side Scatter
TCF4	T-Cell Factor 4
Wnt	Wingless/Int

ACKNOWLEDGEMENTS

Immensurable appreciation and deepest gratitude for the help and support are extended to the following persons who in one way or another have contributed in making this thesis possible.

I care to thank my tutor **Dott. Salvatore**, for the opportunity he gave me of being part of his team at CEINGE, and conceiving my SEMM PhD degree. Without his guidance this study wouldn't have been possible. Thanks to my supervisor **Dott. Del Vecchio**, for his brilliant mentorship and for having introduced me to the bright flow-cytometry universe: no words are able to describe my gratitude. Special thanks to my external supervisor **Dott. Walsh**, for hosting me in his lab at Mount Sinai, and for making me feel part of the "family"; for his great mentorship, advice, support, guidance and suggestions during the last years of my PhD. For his patience and dedication to the guidance of my project during my abroad experience, and especially during the revision and editing of this thesis. For the way he trusted in me since the beginning and for having brought me step by step to my dissertation. Thanks to all my colleagues for their technical support, and for the long scientific conversations and precious advice. I also would like to thant **Valeria Rotoli**, for being the most efficient and professional student's administrator I have ever met; for her tireless and smart personality, and for her essential help and assistance during my entire PhD.

Last but not least, my family for supporting me in every decision and for pushing me when I need.

*My efforts and sacrifices needed for pursuing this work are dedicated to my **Mom** and to her endless strength.*

ABSTRACT

Cancer Initiating cells are a small population of cancer cells capable of tumor initiation and growth. Characteristic of these cells are the expression of "stem-like" markers and the common alteration of the *Wnt*/β-catenin and *Notch* pathways. In my thesis, I characterized Colon Cancer Initiating Cells (CCICs) by FACS analysis from HT-29 colon cancer cell line. I then performed *in vitro* colonospheres-forming assay and RNA-Seq, to interrogate the genome-wide signature involving CCICs. My results show that members of *Wnt/*β-*catenin* pathway are elevated in CCICs and colonospheres, when compared to a more differentiated population. Moreover, genes directing the differentiation were silenced in CCICs and colonospheres. These results demonstrate the self-renewal and proliferation capacity of the isolated CCICs population, and that this subpopulation resembles a strong *Wnt*-signaling pathway signature. Additionally, several lncRNAs are *dys*-regulated in CCICs, and my findings identify *LUST/RBM5-AS1* as a lncRNA transcript strongly elevated during colonospheres formation. The expression of *LUST* and stem-like markers CD24 and CD44, and expression of *Wnt*-signaling corresponds with cells that can survive and grow in serum-free media. Loss of *LUST* impairs *Wnt*-signaling at mRNA and protein levels, while *LUST* overexpression provides enhanced and synergistic signaling mediated through *Wnt* and β-catenin at the mRNA and protein levels. Nuclear/Cytoplasmic RNA fractionation and RNA-FISH show that *LUST* essentially is a nuclear transcript. RNA immuno-precipitation and UV cross-linking immuno-precipitation assays show that *LUST* RNA binds to β-catenin. Finally, *LUST* overexpression enhances colonospheres formation more rapidly. Collectively, my findings reveal that the lncRNA *LUST* regulates *Wnt* pathway in CCICs through a coordinated physical interaction with β-catenin, acting through transcriptional regulation to promote colon cancer initiating cells maintenance.

CHAPTER I: INTRODUCTION

1. Colorectal Cancer

1.1 Colorectal cancer and Intestinal Stem Cells

Colorectal cancer is the third most common cancer and the third cause of cancer death in the United States. In 2014, the estimated new cases for colon cancer and rectal cancer were ~90,000 and 40,000 respectively, with the combined number of deaths annually estimated at ~50,000 (American Cancer Society. Colorectal Cancer Fact and Figures 2014-16). The colon, or large intestine, is anatomically composed of four distinct cellular layers, and the inner luminal layer contains epithelial cells folded into finger-like invaginations, embedded in the sub-mucosal tissue to form the crypt of Lieberkhun [1], functional unit of this organ. In each crypt are present around 2,000 cells, included three mainly differentiated cell lineages: enterocytes, goblet cells and endocrine cells. The differentiated cell lineages reside within the top-third of the crypt [2,3]. These epithelial cells are subjected to a continuous turnover, given that they shed into the lumen once they become senescent [4,5]. Understanding the mechanisms that regulates intestinal stem cells (ISCs) is radical to elucidate the biology of their malignant counterpart. In fact, the ISCs, characterized by self-renewal and multipotent undifferentiated cells, are responsible for the rapid turnover of intestinal epithelium (5 days). Their cell division is mainly asymmetrical: among the two daughter cells generated, one daughter cell is identical to the original cell, while the second forms a progenitor cells with the capacity to differentiate. Progenitor cells migrate to the top of the crypt and reproduce the fully differentiated cell environment [6-8]. The intestinal microenvironment (niche) consists of several components, cellular and extracellular, that provide the best conditions for stem cells maintenance through secretion of growth factors, cytokines and direct interactions [6,9]. Intestinal subepithelial myofibroblasts (ISEMFs) are considered key

regulators of stem cells self-renewal and differentiation and are extremely involved in the interactions between epithelial and mesenchymal cells [3]. These interactions are able to keep the right balance between proliferation and differentiation through the regulation of several pathways as *Wingless/Int (Wnt), Sonic hedgehog (Shh)* and *Notch*.

The cellular process that leads from a normal colonic epithelium to invasive colorectal cancer (CRC) is a multistep oncogenic process, which goes from normal to dysplastic epithelium, formation of adenomatous polyps and, in the later stages, to invasive CRC. Each step in the tumor progression is marked by evolution of prominent genetic alteration. Faeron et al. proposed that mutations in adenomatous polyposis coli gene (APC) arise in the early stages of transformation, leading to hyperproliferation and formation of class I adenoma; class II andenoma are generally induced by *K-Ras* mutations; while loss of 'Deleted in Colorectal Cancer'-DCC- is responsible for class III adenoma; and ultimately, p53 mutations are responsible for invasive cancer [10]. Significant findings have been noted in the frequency of known genetic mutations, like the ones just mentioned, that influence stem cell dynamics in tumor initiation [11]. Furthermore, the extent of stem cell clonal advantage, which constantly increases during tumor progression, is influenced by oncogenic and environmental factors therefore playing an important role in CRC initiation.

1.2 Cancer Initiating Cells (also known as Cancer Stem Cells, CSCs)

Our knowledge of how Cancer Initiating Cells (CICs) promote asymmetrical tumor growth and maintain stem like characteristics is still poorly understood. During the past few decades many new findings have been recognized in understanding the patterns and markers of cells growth within human tumors. However, more detailed investigation is necessary to understand the stem-like character of these cells as initiators of tumor growth and metastasis. A central concept of CICs is that tumors express extensive heterogeneity of cells, with a select few endowed with the capacity to self renew and to maintain this specific function over the evolution of tumor growth and metastasis. It has been carefully documented that cells present in the bulk of a tumor consist of the spectrum of rapidly proliferating cells to postmitotic and differentiated cells. Moreover, the growth of tumors is based on the presence of a small subpopulation of cells with the capability of self-renewal and asymmetrical growth, hence evidence of CICs in CRC [12].

The history of CICs begun with a finding by Furth and Kahn, on 1937 [13], when they demonstrated that a single cells from a mouse tumor, injected into another recipient mouse, could variably initiate a new tumor growth, suggesting the existence of cell-type heterogeneity that defines a small population of cells with a capacity of initiating a new malignancy. Later in the mid-1900s, new approaches became more widely available to identify proliferating cells, as radiolabeling and autoradiography became more prominent [14]. In 1971, using radiolabeled phosphorus on a squamous cell carcinoma, Pierce et al. demonstrated that that most of the undifferentiated areas showed early and rapid labeling of DNA, and, during later stages of cell growth, the well-differentiated areas effectively expressed the DNA radioactive label, demonstrating that the differentiated cells were derived from undifferentiated “stem-like” population of cells. Furthermore, well-differentiated cell population, injected into compatible hosts, was unable to

form tumors [15]. Pierce and colleagues therefore, hypothesized the theory of cancer initiation as: "a concept of neoplasm, based upon developmental and oncological principles, states that carcinomas are caricatures of tissue renewal, in that they are composed of a mixture of malignant stem cells, which have marked capacity for proliferation and limited capacity for differentiation under normal homeostatic conditions, and of the differentiated possibly benign, progeny of these malignant cells".

At the same time the concept of clonal evolution came out from studies performed by Nowell, [16], who proposed that "most neoplasms arise from a single cell of origin, and tumor progression results from acquired genetic variability within the original clone allowing sequential selection, leading to more aggressive subclones. Tumor cell populations are apparently more genetically unstable than normal cells, leading to the greater capacity for cellular heterogeneity. The acquired genetic instability and associated selection process, most readily recognized cytogenetically, results in advanced human malignancies being unique in their karyotype and biological profiles in tumor growth. Therefore, each individual's malignant disease may require personalized approaches to molecular based therapies, and even this may be challenged by the evolution and emergence of a genetically variant subclones resistant to the chemoterapeutic treatment" [16]. About twenty years later, John Dick and colleagues revives CSC theory by xenografting human acute myeloid leukemia (AML), using $CD34^{+}CD38^{-}$ fractions into immunocompromised host mice. In this study the frequency of the initiating cell was found on order of one per million tumor cells, leading to the identification of CIC in AML with unique stem-like character [17,18]. In 2003 Clark and his colleagues [19] applied this system for the first time to a solid breast cancer tumor, and isolated a cell fraction CD44+ CD24-/low that was the only one able to generate new tumors in immunodeficient mice, as well as the nontumorigenic cells present in the initial tumor. Similar studies on brain [20] and colon c

ancer [21-23] were published there after. Therefore, what makes cells CICs is the ability of (1) being the source of all of the malignant cells present in a tumor; (2) resistance to chemotherapeutic agents that makes them responsible for recurrence; and (3) giving rise to metastases.

The combination of several markers can be used to identify and enrich the CICs population in different kind of tumors (**Table 1.1**). Indeed, based on the heterogeneity of several markers, specific subpopulations of cells can be sorted from primary tumors or cell lines and injected in immunodeficient mice by subsequent serial passages. The capacity of this subpopulation of cells for making new tumors is interpreted as evidence of CICs in the primary tumor.

Tumor Type	Cell Surface Markers on CICs	Reference
Leukemia	CD34+,CD38-,HLA-DR,CD71-,CD90-,CD117-,CD123+	Guzman ML, 2004
Melanoma	CD20+	Fang D, 2005
Colon	CD133+,CD44+,CD24+,CD166+,EpCAM+	O'Brien 2004; Dalerba P, 2007; Yeung TM, 2010
Lung	CD133+,ABCG2 (high)	Eramo A, 2008; Ho MM, 2007
Pancreatic	CD133+,CD44+,EpCAM+,CD24+	Li C, 2007; Simeone DM, 2008
Breast	ESA+ CD44+CD24-/(low)	Al-Hajj, 2003

Table 1.1: Cell surface phenotypes of Cancer Stem Cells in different tumor types [24]

1.3 Colon Cancer Initiating Cells, CCICs (also known as Colon Cancer Stem Cells)

CSCs are identified a small subpopulation of cells endowed with the ability of self-renewal, the capacity of giving rise to asymmetrical growth of the parental tumor, possessing intrinsic resistance to chemotherapeutic drugs and elevated insensitivity to radiation, and can form new tumors, metastasize and relapse after tumor regression [25]. The first description of colon cancer stem cells (CCSCs) was provided by Dick and De Maria independently [22,23]. Dick and colleagues demonstrated that by implanting serial dilutions of human colon cancer cells into immunodeficient mice, only a small population equal to $1/5.7 \times 10^4$ of total cells, could induce tumor growth. Using the antibody directed against CD133 antigen (Prominin), they were able to sort two different subpopulations and noticed that the CD133- subpopulation was also able to form tumors but the tumor formation was taking more time compared to the one originated from the CD133+ subpopulation. Dick and colleagues then concluded that CD133- cells were malignant cells derived from asymmetric division of CD133+ cells, and that not every CD133+ cell was able to give rise to tumor formation. Limiting dilution assays showed that only 1 of 262 CD133+ colon cancer cells could induce tumor formation [22]. Ricci-Vitiani group also identified CD133 as a potential CCSCs marker, but they noticed also that the same subpopulation of cells was cytokeratin (CK) 20 negative. In fact, CK20 is a colonic epithelial differentiation marker, so not supposed to be expressed on stem cells. Through *in vitro* colonospheres formation assay, they showed that CD133+ cells could be isolated by plating single cells from colon cancer tissues in serum free media, supplemented with epidermal growth factor and basic fibroblast growth factor. In these conditions, CD133+ cells could form spheres-like aggregates, proliferate at an exponential rate and increase their aggressiveness through many passages *in vivo*. Also, after inducing differentiation by adding serum, they noticed that the expression of CD133 was decreased, along with the gaining of CK20 expression [23]. More evidence by Todaro et al, showed

that CD133+ cells are able to generate colonies organized as a crypt-like structure, when cultured on Matrigel, under differentiation conditions. Furthermore, they demonstrated that this subpopulation of cells was strongly resistant to apoptosis by generating Interelukin-4 (IL-4) [26].

The cell surface protein CD133 is a five-transmembrane glycoprotein identified as the first CCSCs marker [22,23,26,27]. Although several interesting findings were noted with the CD133 cell surface marker, its use as a source of cancer initiating cell populations remains controversial. Many in the field, consider that: (i) cell surface expression of CD133 is not solely restricted to the stem compartment but has been found to be expressed throughout the normal gastro-intestinal tract [28] and (ii): CD133+ as well as CD133- are able to form tumors [29]. This was the first attempt to identify the actual CSCs, but during the years, several additional cell surface antigens have been associated with the cancer stem cell phenotype, i.e.: CD44 [21,30,31] CD166 [21], CD24, CD29 [27], and Lrg-5 [32]. In a hallmark study by Young et al it has been further demonstrated that isolated CD24+/CD44+ subpopulations from different colon cancer cell lines, were capable to self-renew and to reestablish all the CD24/CD44 subpopulations *in vitro* and *in vivo* [33].

In the attempt of finding more specific markers to characterize CCSCs, the cell surface antigens CD24 and CD44 were used in combination with CD133 [22,23,27,34]. CD24 is a small glycoprotein, consisting of 27 amino acids, and is attached to the cell membrane through a phosphatidylinositol anchor [35]. It is a heavily glycosylated protein, expressed in a wide variety of cancer cells and is directly involved in cell-cell and cell-matrix interactions [36-38]. CD24 was first discovered in mice as a heat-stable antigen, expressed on hematopoietic [39] and neuronal cells [40]. At the functional level, CD24 is considered an alternative ligand for P-selectin, the adhesion receptor present on platelets and endothelial cells [41], and their interaction facilitate the passage of cancer cells through the blood stream during formation of metastatic tumors. Therefore is considered to increase

proliferation rate and adhesion of tumor cells to fibronectin, laminin and collagen [42]. Recent studies found that CD24 can be co-expressed with CD29, CD31 and CD44 in various cancers and with different functions [27]. For example, in pancreatic cancer, the expression of CD24 on both, the surface membrane and the intracellular environment, inhibits the cell invasion and metastasis [43], while other studies showed that in lung cancer CD24 is involved in tumor progression [36], and in the formation of metastatic tumors in other solid tumor types [44]. The expression of the cell surface antigens CD24 in combination with CD44 in colorectal cancer stem cells has been shown in several studies [33,45].

The cell surface antigen CD44 belongs to the family of cell adhesion molecules (CAMs). CD44 is a transmembrane glycoprotein and together with selectins, integrins and cadherins controls the interactions between cells or between cells and extracellular matrix, therefore extremely important to sustain the tissue integrity [46]. There are several protein isoforms of CD44, all alternative splicing forms deriving from a single gene present on chromosome 11 in humans, but the smallest and ubiquitously expressed form is the CD44s (Reviewed in [47]). At the functional level, CD44 plays an important role in promoting migration, proliferation, motility, growth, migration, survival, angiogenesis and differentiation [30,48]. It has been shown that CD44 interaction with osteopontin regulates its cellular function and leads to tumor progression [49], it also interacts with P- or L-selectin, and as for CD24, it helps the cancer cells to spread through the blood stream [50]. Interestingly, CD44 expression is regulated by *Wnt* signaling; weather or not the regulation is direct or indirect remains unclear [51], although recent studies showed that CD44 physically associates with LRP6 upon Wnt treatment and modulates LRP6 membrane localization, acting as a positive regulator of the *Wnt* receptor complex [52].

1.4 Wnt/β-catenin signaling

One of the common *dys*-regulated pathways in CRC is the *Wnt*/β-catenin signaling. The *wnt1* gene, initially termed *Int-1,* is a proto-oncogene encoding a cysteine-rich secreted protein, and was first identified in 1982 [53]. A few years later, Nusslein-volhard and Wieschaus discovered the fly wingless (*wg*) gene, which controls segment polarity during larval development and is a homolog of *wnt1* [54]. In early 1990's, epistasis experiments in *Drosophila melanogaster* [55-57] and developmental studies in *Xenopus* [58] demonstrated that the entire *Wnt* pathway was evolutionary and highly conserved and referred to it as the canonical *Wnt* signaling cascade. Major discoveries were noted in late 1990's with the identification of TCF/LEF transcription factors [59,60] and β-catenin [61] as *Wnt* nuclear effectors, Frizzled as *Wnt* receptors [62] and LRPs/*Arrows* as co-receptors [63]. The human genome contains nineteen confirmed *Wnt* genes, falling into twelve conserved *Wnt* gene subfamilies. Of note, most mammalian genomes harbor *Wnt* genes [64] while is interesting noticing how single-cell organism do not have *Wnt* genes, suggesting that this pathway is extremely important in evolution [65].

The *Wnt* family genes are expressed as proteins of approximately 40kDa size, contain many cysteines [66] and are frequently lipid-modified [67] for the induction of signaling and *Wnt* secretion [68,69]. *Wnt* signaling appears to occur mostly between cells in proximity of each other, like the adult stem cell niche [70,71], making the *Wnts* signals acting over a short-range [72] Once the *Wnt* proteins reach the surface of the target cells, they bind an heterodimeric complex formed by a seven-transmembrane (7TM) receptor, named *Frizzled* (*Fz*), that contains a large extracellular N-terminal cystein-rich domain (CRD) [62], that acts as a platform for *Wnt* ligand binding [73,74], and a single-pass transmembrane molecule called LRP5 and-6 in vertebrate [75,76] and *Arrow* gene in *Drosophila* [63]. The altered conformation of the receptor induced by the ligand, is followed by the

Axin binding to the cytoplasmic tail of LRP6 [77], which is phosphorylated by two kinases, glycogen synthase kinase 3-beta (GSK3-β) and casein kinase 1-gamma (CK1-γ). GSK3β is also responsible for the serine phosphorylation in PPPSP motif present in β-catenin, Axin, and APC [78]. At the same time, the cytoplasmic domain of *Fz* interacts with Dishevelled (Dsh) [79] facilitating Axin and LRP6 tail interaction through a common DIX domain [80,81]. The Axin protein acts as scaffold for the assembling of the "destruction complex" consisting of β-catenin, the tumor suppressor APC, and the kinases CDK1-α/γ and GSK3-α/β. APC is able to bind at the same time Axin and β-catenin trough three axin-binding motifs that are interspersed between a series of 15-20 aa repeats that bind β-catenin, and it has been shown that is extremely important for the destruction complex function in colon cancer, although its specific molecular function remains unclear [82]. In an inactive state CK1 and GSK3 sequentially phosphorylate the β-catenin that is bound to Axin, at a series of N-terminal Ser/Thr residues. The phosphorylated β-catenin is then recognized by the specific protein b-TrCP belonging to the E3 ubiquitin ligase complex and targeted for destruction by the proteasome [83]. β-catenin degradation, therefore, prevents the Wnt target genes activation.

In an active state, the complex instead becomes saturated by phosphorylated β-catenin and new protein is synthetized ad accumulated in the cytoplasm, ready for nuclear migration [84]. (**Fig.1.1**)

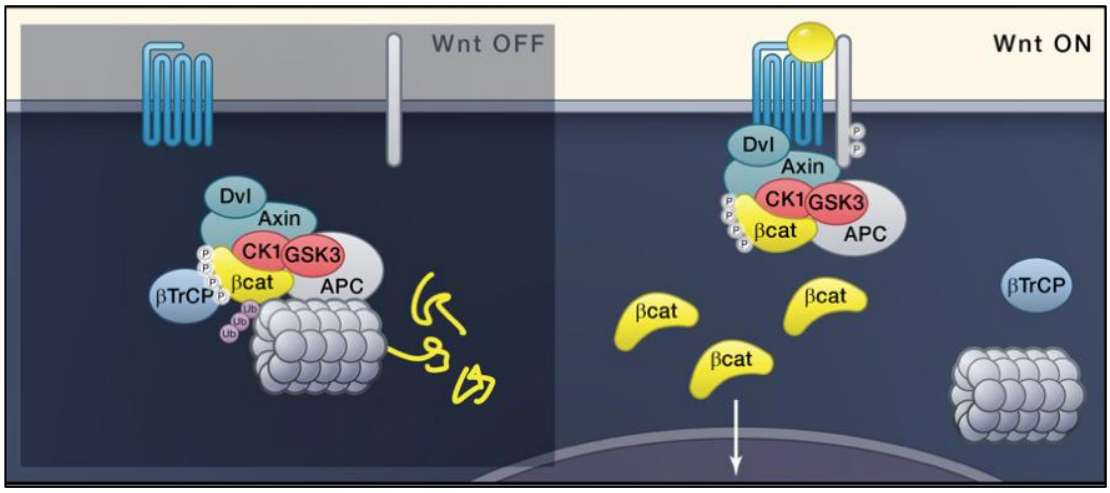

Figure 1.1: Reproduction of *Wnt* pathway in inactive and active state

In the absence of *Wnt* ligand (*left*), the destruction complex resides in the cytoplasm, and induces binding, phosphorylation, and ubiquitinilation of β-catenin by β-TrCP. The proteasome recycles the complex by degrading β-catenin. In the presence of *Wnt* ligand (*right panel*), the intact complex associates with phosphorylated LRP. Once is bound to LRP, the destruction complex stills captures and phosphorylates β-catenin, but ubiquitination by β-TrCP is blocked. Newly synthesized β-catenin accumulates. [72]

Once entering the nucleus, β-catenin engages DNA-bound TCF/LEF transcription factor [59,60]. When the system is turned off, TCF interacts with a member of the *Groucho* (also referred to as the Transducin-like enhancer) family of transcriptional repressors, and the gene transcription is repressed. In the active state, TCF/LEF and β-catenin associate and TCF/LEF is transiently converted in a transcriptional activator of *Wnt*-mediated target genes [85,86] as *MYC* [87] and *CCND1* [88], involved in cell proliferation. (**Fig. 2.1**)

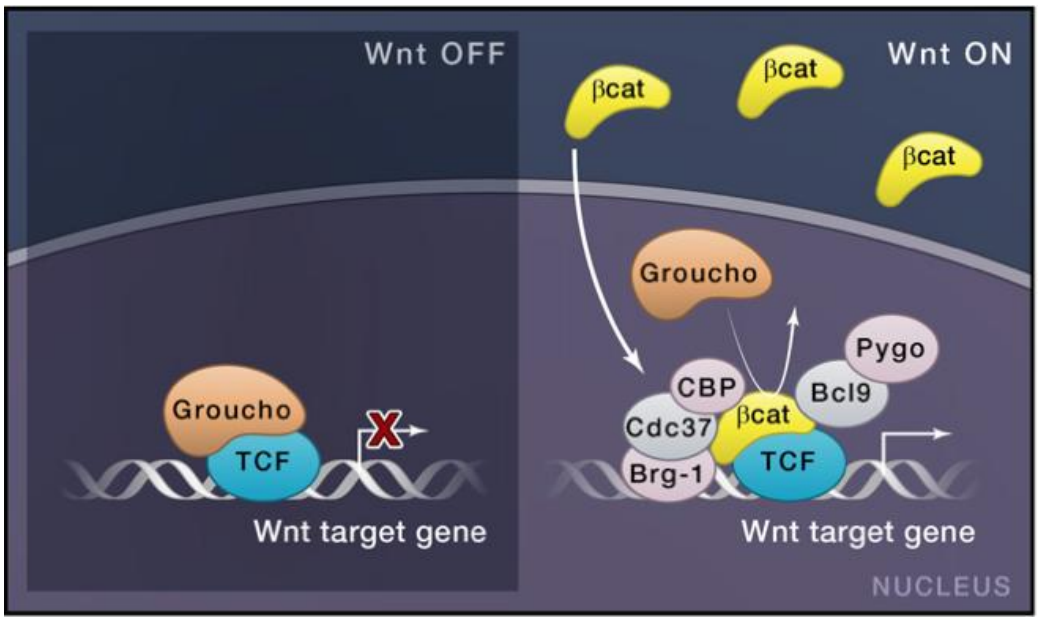

Figure 2.1: Nuclear Wnt-signaling

In absence of *Wnt* ligand, TCF and the transcriptional corepressor Groucho, occupy and repress *Wnt*-target genes. When *Wnt* signal is present, β-catenin replaces Groucho from TCF and recruits other transcriptional activators and histone modifiers such as Brg1, CBP and other. [72]

Although the mechanism of how β-catenin shuttles between the cytoplasm and the nucleus remains to be clearly resolved, it's been shown that the fold change rather than absolute levels are critical, suggesting that low levels of this protein is capable of inducing important transcriptional changes [89].

2. Long Noncoding RNAs

From a pure molecular perspective, cancer is a disease involving both, genetic and epigenetic mechanisms, due to aberrant function and expression of genes regulating tumor suppression and oncogenesis and at the same time, hundreds to thousands of lncRNAs have been involved in cancer and oncogenesis. In the following paragraphs, I provide a comprehensive introduction from existing literature of lncRNAs and their emerging roles in cancer and cancer initiating cells related to the findings in the thesis.

2.1 Biogenesis and function of lncRNAs

New techniques for genome sequencing have shown that the majority of the eukaryotic genome is pervasively transcribed, but less than 2% is ultimately translated into proteins [90-92], suggesting that non-coding RNAs impose another layer of regulation functioning throughout the human genome. Non-coding RNAs (ncRNAs) can be distinguished in different categories according to their length, expression and function. The small ncRNAs class includes microRNAs (miRNAs), small interfering RNAs (siRNAs), and Piwi-associated RNAs (piRNAs), all of them shorter that 200 nucleotides [93]; RNA transcripts with a length above 200 nucleotides, often polyadenylated and devoid of evident open reading frames (ORFs), are defined as long noncoding RNAs (lncRNAs) [94-96].

According to the accumulated datasets from the ENCODE Project Consortium as of the year 2012, there are approximately 9640 lncRNA loci in the human genome that have been recognized [95,97], and as sequencing technology rapidly emerges from single cell approaches to increases in the depth of coverage, lncRNA populations are rapidly being identified. Several studies showed

that they are involved in the regulation of chromatin organization, transcriptional and post-trascriptional levels [98], and in tumorigenesis [99].

Based on LNCipedia 2.0, the lncRNAs database (http://lncipedia.org/db/search) there are 32,183 human annotated lncRNAs, and among them only very few have been validated either functionally or biologically [100]. The transcription of lncRNAs can arise from intergenic or intragenic regions; the lncRNAs arising from intergenic regions are called large intergenic ncRNAs (lincRNAs), and are typically located 5kb away from protein coding genes [101]. LncRNAs arising from intragenic regions can be further classified as in antisense, overlapping, intronic, with bidirectional orientations relative to protein-coding genes or gene regulatory regions, such as UTRs, promoters and enhancers (Reviewed in [102]). Like their protein-coding gene counterparts, ncRNA transcript expression can be controlled by transcriptional and epigenetic factors, in particular multiple transcription factors (TFs) such as Nanog, Oct4, NFkappaB, Sox2 and tumor protein 53 (TP53) are found to transactivate their expression [101,103]. Again, like their protein-coding gene counterparts, lncRNAs are subjected to post-transcriptional processing and modification [95,97] such as 5' capping, polyadenylation, alternative splicing or RNA editing [104]. At the structural level, given their length, lncRNAs maintain complex secondary and tertiary structure, favoring the duplexing with DNA and other RNA molecules as well as through interactions with numerous RNA binding proteins, to impose biochemical and biological activities. Therefore, lncRNAs can function as activators, decoys, guides, or scaffolds for transcription factors and histone modifiers [105] (**Fig. 3.1**) as described below:

1. Long non-coding RNAs can function as transcriptional activators by directly interacting with transcriptional factors, favoring an allosteric change to promote trans-activation of target gene transcription (**Fig. 3.1**, *top left*).
2. Long non-coding RNAs can silence transcription factor function by sequestering transcription factors away from chromatin (**Fig. 3.1**, *top right*).
3. Long non-coding RNAs can function as transcriptional guides by recruiting chromatin-modifying enzymes to target genes, either in *cis* or in *trans* and altering gene expression (**Fig. 3.1**, *bottom left).*
4. Long non-coding RNAs can function as scaffolds (platforms) for chromatin modification complex that form ribonucleoprotein complexes (**Fig. 3.1**, *bottom right).*

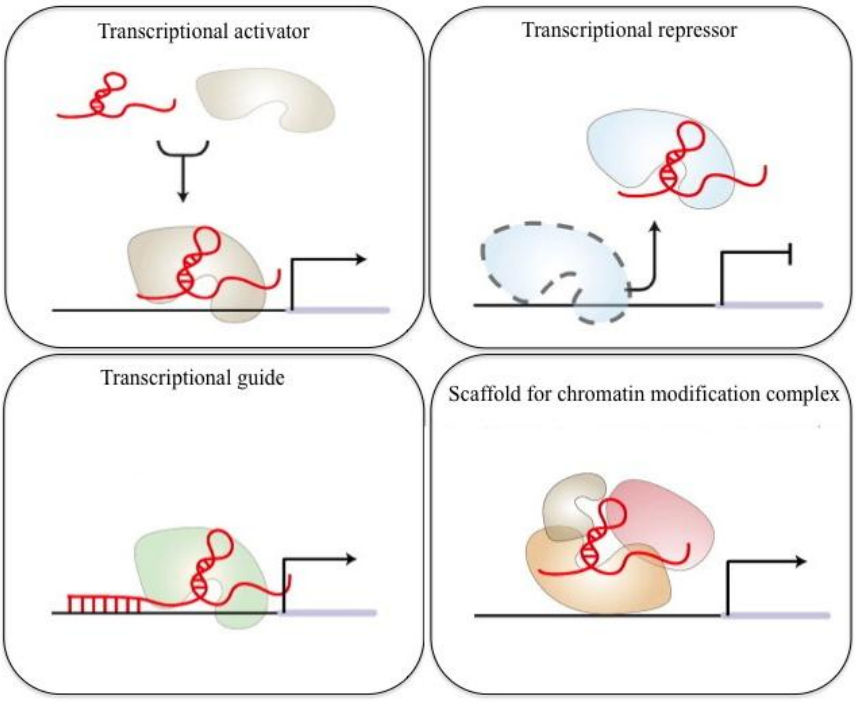

Figure 3.1: Functions of lncRNAs

Top left) transcriptional activators; *top right*) transcriptional repressor; *bottom left*) transcriptional guide; *bottom right*) scaffold for chromatin modification complex. (Modified from [106])

Examples of lncRNAs involved in some of these complicated interactions are reported in **Table 2.1.**

Interaction	lncRNA	Function	Ref
lncRNA:Chromatin regulators	HOTAIR:LSD1-CoREST Xist: PRC2	Targets the LSD1 complex to demethylate H3K4me2 Targets PRC2 to mediate H3K27 methylation	Tsai MC.2010 Zhao J.2008
lncRNA:TFs	Gas5:Glucocorticoid rec PANDA:NF-YA	Titrates GR away from target genes Titrates NF-YA away from apoptotic genes	Kino T.2010 Hung T.2011
lncRNA:mRNA	LincRNA-p21:JUNB	Induces degradation of target mRNA	Yoon JH.2012
lncRNA (miRNA)	H19 (miR-675)	Produces miR-675	Steck,2012
lncRNA:RNA binding protein	LincRNA-p21:HuR Malat1:Ser/Arg (SR) prot	Induces degradation of lincRNA-p21 Modulates SR splicing factors	Yoon JH.2012 Tripathi V, 2010

Table 2.1: Example of different kinds of lncRNAs interactions (Modified from [106])

2.2 lncRNAs in cancer

2.2.1 Genetic and epigenetic changes deregulating lncRNAs

Previous known variation in DNA copy number (or copy number variation) commonly occur in non-coding regions of the human genome, and had remained inexplicable. Because of recent advancements in genome–wide sequencing technologies, many of these same regions were revealed as transcribed elements and can involve microRNAs, SINE, LINE transcripts and lncRNAs like *LOC285194* which has been demonstrated to act as a tumor suppressor. Furthermore, genetic deletion of the *LOC285194* locus is associated with poor survival of patients affected by osteosarcoma [107]. Loss of heterozygosity (LOH) of the maternal allele of *KCNQ1OT1* is also associated with cancer, as well as gene deletion of *PTENP* (*PTEN*-Pseudogene) in melanoma patients [108].

A study conducted by Akrami et al. identified that an intergenic lncRNA on chromosome 1, *OVAL*, shows narrow focal genomic amplification in a defined subset of ovarian cancer, providing evidence of the strong effect of lncRNAs in cancer development [109]. The emerging role of the genetic changes affecting the lncRNAs expression has been shown in several types of cancer. For example, the effect of the single nucleotide polymorphisms (SNPs) rs1456315 and rs7463708 is associated with cancer susceptibility and expression of the oncogenic lncRNA *PRNCR1* in prostate cancer [110].

Several additional lncRNAs are localized within the imprinted loci, playing an important role in the imprinted loci formation and maintenance [111], as their localization make loci being susceptible to epigenetic activation or inactivation. Few examples are represented by *H19* that is de-repressed by Mineral dust-induced gene (*Mdig*). The *Mdig* gene is responsible for the down-regulation of H3K9me3 and heterochromatin in the *H19* loci, inducing the increase of

the oncogenic H19 [112]. Defect of imprinting correlating an increase of lncRNAs expression are also reflected by *KCNQ1OT1* in multiple cancers [113-116].

2.2.2 LncRNAs in cancer initiation, progression and metastasis

Tumor formation is the result of the interaction between the cancer cell and the microenvironment. Dependent on these intricate interactions, cancer cells are able to acquire the ability for sustaining proliferative signaling, enabling replicative immortality, inducing angiogenesis, evading growth suppression, resisting cell death, activating invasion and even acquire drug resistance [117]. The three stages of cancer are characterized by initiation, growth and metastasis, and a large number of protein coding genes have been demonstrated to be involved in these processes. Despite this knowledge, how these genes are deregulated is not fully understood. Little is known about noncoding genes and their role in cancer, but emerging evidence show that noncoding RNAs play a role in cancer development.

Expanding findings, based on genome-wide sequence information and animal modeling using selected cell populations from tumors, suggest that specific cells acquire traits similar to embryonic stem cells for asymmetrical growth and cellular diversification. Through this evidence, key genes and proteins governing the embryonic and adult stem cell multipotency are identical or similar in cancer stem cells [118]. Regarding cancer initiation, it has been demonstrated from Trimarchi et al., that there is a set of T cell acute lymphoblastic leukemia (T-ALL) associated lncRNAs directly regulated by the Notch/Rpbjk activator complex. In particular, the lncRNA *LUNAR1*, is required for T-ALL growth *in vivo* and *in vitro*, given to its potential of inducing *IGFR1* mRNA and therefore sustaining IGF1 signaling [119]

.

The *Polycomb* repressive complex-1 (PRC1) and -2 (PRC2) promote stem cell self-renewal and inhibit differentiation, playing a role in tumor progression and development. PRC1 core complex includes BMI1, mPh1/2, Pc/Chromobox (CBX) and the ubiquitin E3 ligase Ring A/B proteins, while EED, SUZ12 and EZH 1/2 are mainly members of PRC2 complex [120]. A growing body of evidence shows that members of the *Polycomb* repressive complex interact with lncRNAs [121]. The lncRNA *ANRIL* physically interacts with CBX7, a member of the PRC1 complex, where this interaction forms heterochromatin surrounding the *INK4b-ARF-INK4a* locus, leading to its repression. This mechanism facilitates the bypassing of senescence, endowing stemness of PC3 prostate cancer cells for sel-renewal [122,123]. Another lncRNA involved in PRC interaction is the oncogenic *HOTAIR*, responsible for the relocalization of the PRC2 complex towards a metastasis expression profile [124]. Other studies include lncRNAs lined with cell cycle pluripotency regulators, such as *p53, Pou5f1/Oct4, NFkappaB* and *Nanog* [101,125], essential for eradication or establishment of cancer stemness respectively [126].

The process that leads to the formation of a visible tumor mass, arising from a few cancer initiating cells, is defined as cancer growth referring to tumor cell proliferation, resistance to death (apoptosis) and angiogenesis. Several lncRNAs have been identified as contributor to cancer cell proliferation, and an example is given from *PCAT-1,* which can promote prostate cancer cell proliferation [127] as well as *PCGEM1* [128-130]. The lncRNA *H19* has been characterized for its ability to promote the anchorage-independent growth of primary breast and lung carcinoma cell lines [131,132]. On the other hand, the growth arrest-specific 5 (*Gas5*) lncRNA, associated with the ability of inducing cell growth arrest, was found down-regulated in breast cancer [133].

Angiogenesis occurs upon hypoxia and becomes another acquired characteristic of the tumor. Several studies identified lncRNAs involved in angiogenesis, and in particular, genetic deletions or pharmacological inhibition of *MALAT1* reduce vascular growth *in vivo*, as shown from Michalik et al [134]. Furthermore, elevated *MALAT1* transcript expression was found in colon, lung and liver cancer [135-137], suggesting a specific role in cancer progression by promoting angiogenesis. Other studies suggest that *MEG3* may inhibit angiogenesis in vivo [138] and LOC285194 could inhibit the VEGFR1 [107]. Following angiogenesis, cells adopt a migratory mesenchymal character to facilitate colonization in different tissue through a peripheral blood vessel, thereby forming metastasis. Some lncRNAs have also been directly lined to the metastatic process. For instance, *HOTAIR* overexpression leads to breast cancer metastasis as seen in an *in vivo* assay, by altering the transcription profile and favoring de-differentiation of the subcellular populations into mesenchymal–like character followed by metastasis. HOTAIR lncRNA expression is increased in 25% of breast cancer patients, whose tumor staging correspond with the appearance of distant metastatic tumors [124]. High expression levels of this lncRNA, has also been observed in advanced stages of colon, liver and pancreatic cancer [139-142]. The lncRNA-activated by TGF-b (*lncRNA-ATB*) was found to be upregulated in hepatocellular carcinoma metastases and has been associated with poor prognosis. Expression of *lncRNA-ATB* correspond with ZEB1 and ZEB2 expression by competitively binding the mir-200 family, thereby inducing EMT and invasion. In addition, the same lncRNA promotes organ colonization of disseminated tumor cells by binding *IL-11* mRNA and triggering STAT3 signaling [143].

2.2.3 LncRNAs as diagnosis markers and therapeutic targets

Analogous to the detection of miRNAs, lncRNAs can be detected from various body fluids by RT-PCR and suggest their putative use as diagnostic and prognostic indicators of disease outcomes. Although these studies remain underdeveloped at the current time, emerging technologies show how is it feasible to detect the lncRNA *PCA3* in prostate cancer patient urine samples, or the lncRNA transcript *FR0348383* in post digital rectal examination (DRE) urine [144], and the lncRNA *HULC* in HCC patient blood [145-147] as transcript signatures of malignant disease. A recent study from Tong et al. showed that the lncRNA *POU3F3* can be detected from esophageal squamous cell carcinoma (ESCC) patient plasma samples, as a potential biomarker for early screening of ESCC [148].

2.3 The lncRNA LUST (RBM5-AS1)

In 2009, Rintala-Maki et al identified and characterized a novel antisense non-coding RNA from the RNA binding motive 5 (RBM5) locus [149]. The RBM5/LUCA-15/H37 gene is transcribed in both the sense and antisense orientations [150,151] and encodes products that are able to differentially modulate apoptosis, depending on the transcript splice variant that arises. Alternative splicing of *RBM5* pre-mRNA produces at least five variants: RBM5, RBM5Δ6, RBM5+5+6, RBM5+6 and RBM5+5+6t/Clone 26 [152]. In overexpression studies from various cell line, the full length *RBM5* transcript sensitizes cells to apoptotic signal and varies with cell type [153-157]. The alternative splice variant RBM5Δ6, lacking exon 6, inhibits apoptosis and increase cell growth when overexpressed in the CEM-C7 leukemia cell line [155]; RBM5+6, including intron 6, and RBM5+5+6, including both intron 5 and 6, contain a premature stop codon, thus targeting the transcripts as candidates for nonsense-mediated decay (NMD). RBM5+5+6t/Clone 26 (GenBank accession no. AF107493) is a truncated form of RBM5+5+6 and is pro-apoptotic/cytotoxic; the small cDNA Je2 (326 bp) was originally identified by Sutherland et al. as an antisense to RBM5+5+6t/Clone 26 and they demonstrated that is able to confer dramatic apoptosis resistance to cells treated with normally lethal levels of agonistic anti-Fas antibody [150]. Several years later, the same group discovered that Je2 was only a fragment of a larger transcript, and in particular of a transcript of ~1.4kb length termed *LUST* (Luca-15-specific transcript). They showed that *LUST* is an antisense RNA, is a ncRNA and regulates RBM5 variant expression, and speculated that *LUST* functions to enhance full length RBM5+5+6 expression, which is then directed to the NMD, and suppresses RBM5+5+6t/Clone 26, thereby inhibiting cell death. In their model, they specifically proposed that the antisense ncRNA *LUST* forms a double-stranded RNA duplex with the alternatively spliced intron 5 and 6 retaining RBM5 transcript, soon after it emerges from the RNA

polymerase II transcription elongation complex [149]. However, the biological function of LUST remains poorly studied.

CHAPTER II: MATERIALS AND METHODS

Cell culture

Human LS174T (ATCC # CL-188) and SW480 (ATCC #CCL-228) colon cancer cell lines were obtained from CEINGE-Advanced Biotechnology (Naples, Italy) cell culture facility, where were maintained according to the original datasheet. Human HT-29 (ATCC #HTB-38) and CaCo2 (ATCC #HTB-37) colon cancer cell lines were purchased from ATCC and tested for mycoplasma contamination using LookOut Mycoplasma PCR Detection Kit (Sigma-Aldrich) according to the standard protocol. HT-29 cells were grown in Dulbecco's modified Eagle medium (DMEM) supplemented with 10% Fetal Bovine Serum (FBS) (Cornig), 1% penicillin-streptomycin (Cornig) and 2% Glutamine (Cornig). CaCo2 cells were grown in DMEM medium supplemented with 20% FBS (Cornig) and 1% penicillin-streptomycin (Cornig). Both cell lines were grown in adherent conditions by standard methods at 37°C in 5% CO_2. The medium was changed twice a week, cells were passaged using 0.05% trypsin/EDTA (Cornig) and used at early passages.

Flow-cytometry analysis and cell sorting

Multi-color flow-cytometry was performed with anti-human monoclonal antibodies (MoAbs) that were conjugated with fluorescein isothiocyanate (FITC), phycoerythin (PE), allophycocyanin (APC) and phycoerythin-Cy7 (PE-Cy7). Cells were stained using FITC-conjugated MoAbs against CD24, PE-coniugated MoAbs against CD166, APC-coniugated MoAbs against CD133 (BD Biosciences), and PE-Cy7-conjugated MoAb against CD44 (BioLegend). An analysis buffer composed of RPMI without red phenol (Invitrogen), 1-2% FBS (Cornig) and 10U/ml DNase (Sigma-Aldrich) was used to prepare cells for the analysis.

Enzymatically individualized cells were counted, resuspended in analysis buffer at 5 x 10^6 cells/ml and stained by incubation at 4°C for 20 min with the appropriate MoAbs. For FACS analysis, cells were stained in a 100µl labeling volume using the following concentrations, respectively: 0.1:10 for CD44 and 0.5:10 for CD24, CD133 and CD166. Samples were washed twice with analysis buffer, centrifuged and resuspended in 0.5ml of FACS flow sheath fluid (BD Bioscience). Immediately before FACS acquisition, cells were incubated at room temperature in the dark with DAPI (Invitrogen) to exclude dead cells. All experiments included a negative control to exclude the signal background caused by cellular auto-fluorescence.

Fluorescence-activated cell sorting of HT29 cells was performed after cells were enzymatically individualized, resuspended at 5 x 10^7 cells/ml in sorting buffer (RPMI without red phenol (Invitrogen), 1-2% FBS, 10U/ml DNase, 2.5mM EDTA)) and stained by incubation with CD24 and CD44 MoAbs at 4°C for 20 min. After the staining, cells were washed twice with sorting buffer and resuspended at 2 x 10^7 cells/ml. The samples were sequentially filtered (50µm, Partech) and incubated for 5 min in the dark with a vital dye, DAPI (Invitrogen).

Samples were analyzed on a BD LSRII four lasers flow-cytometer (Bekton Dikinson, Franklin Lakes, NJ, USA). For each sample run, 10^4 to 2 x 10^4 events were recorded and analyzed. HT-29 live cells sorting experiments were performed using BD FACSAria III using a 70µm nozzle. Sorted cells were collected in DMEM medium supplemented with 10% FBS, 1% penicillin/streptomycin and 2% Glutamine.

Analysis of cytometric data was performed using FACSDiva software (Bekton Dikinson). A three-gating strategy to define the target cell population was defined. Cells were initially gated

on a two physical parameters dot plot measuring forward scatter (FSC) *versus* side scatter (SSC) to exclude most of the dead cells and debris. Second, doublets were excluded by gating cells on FSC-height *versus* FSC-area dot plots. The last gate was used to exclude dead cells, based on DAPI expression. Finally, cells were gating according to their Bright or Dim co-expression of CD24 and CD44.

In vitro Colonospheres formation assay

The ability of the cells of forming colonospheres is strictly dependent on their stemness potential. In particular, cells enriched on stemness markers are able to grow in serum-free media, in a non-adherent manner and are able to generate colonospheres [22,23,26]. Colonospheres formation assay was performed as previously described [158]. Briefly, cells originating from the total HT-29 colon cancer cell line, or HT-29 sorted cells were plated at a density of 30,000 cells/ml in stem cell medium, using low attachment plates (Corning). Stem cell medium (SCM) was composed of 1:1 ratio of DMEM/F12 (Gibco), 1% penicillin/streptomycin (Cornig), 2ml 50X B27 supplement (Invitrogen), 4µg/ml heparin (Fisher), 1% (w/v) of non-essential amino acids (Cornig), 1% (w/v) of sodium pyruvate (Cornig), 1% (w/v) of L-glutamine (Cornig), 10ng/ml FGF, 20ng/ml EGF, 1X growth factor mix (Growth factor mix 10X: 100mL DMEM/F12, 4ml 30% (w/v) glucose, 200mg transferrin, 50mg insulin in 18ml of water and 2ml of 0.1N HCl, 19.33mg putrescine (sigma) in 20ml water, 200µl 0.3M sodium selenite, 20µl 2mM progesterone (Sigma), H_2O to 200ml). Cells were passaged every 7 days through enzymatic sphere disruption. Briefly, spheres were collected, gently centrifuged for 3 min at 1300 r.p.m. and then washed with PBS. The pellet was resuspended in 1ml of 1X trypsin/EDTA and pipet up and down for 3 minutes, before incubating the cells for 3 min at 37°C and 5% CO_2. After incubation, an equal volume of SCM was added to quench the

reaction. Cells were subjected to a second centrifugation to eliminate residual trypsin and then resuspended in 1 ml of SCM. Trypan blue was used to count the number of viable cells. Cells were passaged for 7 weeks and then induced to differentiate by using DMEM supplemented with 10% FBS, 1% penicillin-streptomycin, and 2% Glutamine.

RNA extraction, cDNA synthesis and qRT-PCR

Total RNA was extracted from HT29, CaCo2 and HT29 derived colonospheres using Trizol and the RNeasy MiniKit (Qiagen) according to the manufacturer's protocol. One to five µg total RNA was reverse transcribed using PrimeScript RT Reagent kit (Takara #6130). Briefly, the RNA was incubated for 5 min at 65°C, in a final volume of 10µl with 10mM dNTP mixture and 50µM of Random 6mers. Subsequently, the template RNA Primer Mixture was mixed with 5X PrimeScript Buffer, 40U/µl RNase Inhibitor and 200U/µl PrimeScript RTase in a final volume of 20µl. The reaction mixture was incubated immediately under the following conditions: 30°C for 10 min, 42 °C for 60 min and inactivation step at 70°C for 15 min. Alternatively, the SuperScript® VILO™ cDNA Synthesis Kit (Life Technologies) was used for low amount of RNA input. Quantitative PCR (qRT-PCT) was performed using the GoTaq® qPCR Master Mix (Promega) on the Stratagene Mx3005P Real-Time PCR System (Agilent Technologies). Hypoxanthine-guanine phosphoribosyltransferase (*HPRT*) gene was used as housekeeping gene for normalization. Sequences of the primers used for qRT-PCR are listed in **Table 1.2.**

Gene	Forward	Reverse
AXIN-2	TTCTGCAAAAGAGAGCTTCCA	ATTGCATCCGTTGCATTCTC
β-catenin	GTGGGACACAGCAGCAATTT	CATCCCGAGCTAGGATGTGA
CCND1	AGAGGCGGAGGAGAACAAAC	GGCGGATTGGAAATGAACTT
CD24	CAGCCAGTCTCTTCGTGGTC	TCCTTGCCACATTGGACTTC
CD44	CAACTCCATCTGTGCAGCAA	CATTGGGCAGGTCTGTGACT
LRG4	TAAAGGAAGGCACCTTTCAA	AGTTATTGGCCCAAGTGTGG
LUST	GCTTCAACACTGCGTGACAA	CGTGGAATCAAATGGAGTGG
MALAT1	GACGGAGGTTGAGATGAAGC	ATTCGGGGCTCTGTAGTCCT
MYC	GGGCTTCTCAGAGGCTTGG	GTCCTTGCTCGGGTGTTGTA
TCF-4	GGAAACCCACCTCCACACTT	GCCAGGCGATAGTGGGTAAT

Table 1.2: List of primers used for qRT-PCR

RNA-seq library preparation

RNA-seq library preparation was performed at the Weill Cornell Medical College Genomic Core facility (New York) using the TrueSeq RNA sample preparation kit (Illumina RS-122-2001) as per manufacturer recommendations. Samples were sequenced by the Illumina HiSeq 2500 platform (Illumina) as 100 bp pair-ended reads.

RNA-seq and Gene Ontology (GO) analysis

Contaminant (aligned) RNA-seq reads were filtered with Bowtie software [159] and aligned to several human reference databases including the human genome (hg19, Genome Reference Consortium GRCh38), RefSeq exons and splicing junctions using BWA alignment algorithm (http://bio-bwa.sourceforge.net/) [160]. The reads that were uniquely aligned to the exon and splicing-junction sites for each transcript were then counted as expression level for a corresponding transcript and were subjected to log2 transformation and global median

normalization. Differentially expressed genes were identified by the R package DEGseq [161] using a false discovery rate (FDR) < 0.001 and fold-change >1.5. Gene ontology analysis was performed using the web tool The Database for Annotation, Visualization and Integrated Discovery (DAVID) (http://david.abcc.ncifcrf.gov/) [162,163].

Human LncProfiler qPCR array

The LncProfiler^{TM} Kit (System Biosciences) is produced to tag and convert small non-coding RNAs into detectable and quantifiable cDNAs. The kit includes assays in pre-formatted plates for 90 well-annotated human lncRNAs, with 5 endogenous reference RNA controls as normalization signals. All of the lncRNAs on the qPCR array have validated primer sets for well-annotated lncRNAs that are registered in the lncRNA database created by Dr. John Mattick (www.lncrnadb.org). Briefly, 1µg of total RNA extracted from $CD24^{Bright}/CD44^{Bright}$ and $CD24^{Dim}/CD44^{Dim}$ HT-29 sorted cells was diluted to 200~400ng/µl. 5µl Total RNA were combined with 2µl 5X PolyA Buffer, 1µl 25mM $MnCl_2$, 1.5µl 5mM ATP and 0.5µl PolyA Polymerase in a final volume of 10µl. The reaction mix was then incubated for 30 min. at 37°C. After the PolyA Tail step, 0.5µl of Oligo dT Adapter were added to the mix to induce the annealing of anchor dT adapter, heating the mix for 5 min. at 60 °C. Once the mix was cooled down at room temperature for 2 min., 4µl 5X RT Buffer, 2µl dNTP mix, 1.5µl 0.1M DTT, 1.5µl Random Primer mix and 1µl of Reverse Transcriptase were added to the mix to induce cDNA synthesis, and the reaction was incubated for 60 min at 42°C and heat for 10 min at 95°C. Following the cDNA synthesis, the qRT-PCR reaction was set up using 1,750µl 2X SYBR Green, 20µl LncRNA cDNA (from the previous step) and 1,730µl of RNase-free water in a total volume of 3,500. 28µl of Mastermix and 2µl of LncRNA primers per well were aliquoted in the qPCR plate. The qPCR reaction was set up using the fallowing steps:

1. 50°C 2 min.
2. 95°C 10 min.
3. 95°C 15 sec.
4. 60°C 1 min.

(40 cycles of Stage 3)

A Dissociation Stage after the qPCR run was added to assess the Tm of the PCR amplicons to verify the specificity of the amplification reaction.

qRT-PCR results were analyzed and ΔΔCT was evaluated using the array related software www.systembio.com/LncRNA.

LncRNA *LUST* knockdown

LncRNA *LUST* knockdown was performed using LNA longRNA GapmeR *in vitro* standard (Exiqon, #300600). LNA^{TM} longRNA GapmeRs are antisense oligonucleotides with perfect sequence complementary to their target RNA. When introduced into the cells, they sequester their target RNA in highly stable DNA:RNA heteroduplexes, leading to RNase H mediated target degradation. The sequences of the oligonucleotides and their LNA^{TM} spiking patterns were carefully designed by Exiqon's GapmeR Design Algorithm (http://www.exiqon.com/ls/Pages/GDTSequenceInput.aspx?SkipCheck=true) to achieve high target affinity with excellent sequence specificity and biological stability. Four different probes directed against lncRNA *LUST* transcript and one unspecific Negative control probe were used.

LncRNA *LUST* cloning and overexpression

The construct pcDNA3-*LUST* was made as previously described [149] with some modifications: *LUST* was amplified from 100ng of HT-29 DNA by PCR using 0.5μM of FactorXFBamHI (5′-CGGGATCCAAATGCCGCCACAGACTTTCA-3′) and 0.5μM of FactorXRNotI (5′-ATGCGGCCGCCAGAAGAATCGCTTGAATCC-3′) primers. The PCR reaction was performed according to the fallowing conditions: 95°C for 5 min., 35 cycles at 95°C for 30 sec., 55°C for 45 sec., 72°C for 80 sec., and a final step at 72°C for 5 min. The amplified product was visualized on 1.5% agarose gel and the corresponding 1.3 kb band was cut and purified. The product was digested with BamHI and NotI restriction endonucleases (New England Biolabs) at 37°C for 1h and purified. Subsequently, the insert was ligated into the pcDNA3 expression vector (Invitrogen) using T4 DNA Ligase (Invitrogen); the ligation reaction was performed for 15 min. at room temperature. Competent cells were transformed with 1μl of ligated vector pcDNA3-*LUST* and grown overnight at 37°C on plates coated with agar, supplemented with 1% ampicillin. A second batch of competent cells was transformed using the digested vector as control. The day after, 12 different colonies from pcDNA3-*LUST* agar plate, were picked and grown overnight at 37°C, 250 r.p.m. Plasmid DNA was extracted using NucleoSpin Plasmid kit (Macherey-Nagel) and submitted for sequencing.

TOPFlash dual luciferase assay

HT-29 cells were seeded at a density of 0.25 x 10^5 cells/well in 12-well plates one day before transfection. Cells were transiently transfected using Lipofectamine (Invitrogen) according to the manufacturer's instructions, using 250ng of the TOPFlash reporter gene construct (M50 Super 8x TOPFlash, Plasmid #12456, Addgene) and 500ng of pcDNA3-LUST and/or 500ng of pcDNA-β-Catenin construct. 48h after transfection, cells were collected and lysed for 15 min. at room temperature. 20µl of each lysate was used to measure luciferase reporter gene expression (Dual-luciferase Reporter assay System, Promega). The luciferase activity was normalized to Renilla luciferase activity from co-transfected internal control plasmid pRL-CMV.

Western blot

For western blot analysis, 30µg of protein lysate were analyzed by SDS-PAGE, transferred to PVDF membranes (Bio-Rad) and blotted with indicated antibodies followed by ECL detection (Thermo Scientific). Western blot assays were performed using the fallowing commercially available antibodies, at the indicated concentrations: anti–β-actin (Sigma, A5441, 1:1,000), anti-α-tubulin (Sigma, T5168, 1:1,000), anti-β-catenin (Bethyl Laboratories, A302-012A, 1:1,000), anti-active-β-catenin (Millipore, 05-665, 1:1,000), anti-cyclin D1 (CCND1, Abcam, ab16663, 1:1,000), anti-c-myc (Cell Signaling, 5605, 1:1,000).

Tansfections

For lncRNA *LUST* knockdown, 1.8 x10^5 HT-29 cells/well were seeded in a 6-well plate and transfected using 300pmole of LNA GapmeRs and Lipomefctamine 2000 according to the manufacturer's protocol. After 48 h, cells were harvested and RNA and protein were collected for qRT-PCR and western blot assays.

For lncRNA *LUST* overexpression, pcDNA3 and pcDNA3-*LUST* construct were transiently transfected into HT-29 cells grown at 30-50% confluence in 100mm^2 culture dishes, washed with OPTI-MEM media (Invitrogen). 4µg of DNA, mixed with 20µl PLUS Reagent (Invitrogen), 20µl of Lipofectamine Reagent (Invitrogen) and 6.5 ml of OPTI-MEM media were added to the cells. Transfectants were incubated at 37°C for 18h, after which media was replaced with DMEM supplemented with 10%FBS, 1% pennicilin/streptomycin and 2% Glutamine. RNA and proteins were collected 24h after transfection to perform qRT-PCR and western blot, respectively.

For colonospheres formation assay, cells transiently transfected with pcDNA-*LUST* were trypsinized and collected 24h after transfection and seeded in a low-attachment plate at 30,000 cells/ml according to the methods previously described.

Nuclear/Cytoplasmic RNA fractionation

For nuclear/cytoplasmic RNA fractionation, HT29 and CaCo2 cells were washed twice and harvested in PBS. After centrifugation at 4°C for 5 min. at 1400 r.p.m, the supernatant was removed and the pellet resuspended in 500µl of RNA Lysis Buffer 1X (2X Lysis Buffer: 280mM NaCl, 3mM $MgCl_2$, 20mM TrisHCl pH 7.5, 0.1% NP40 (igepal)). An equal volume of RNA Lysis Buffer/Saccarose (2X RNA Lysis Buffer, Sucrose 50%) was added at the bottom of a 1.5 ml eppendorf tube to achieve two separate phases, and combined with the cell lysate. After centrifuging at 4°C for 10 min. at 13,000 r.p.m., the nuclei were obtained by deposition at the bottom of the tube, and the cytoplasm at the top. 450µl of the cytosolic fraction were collected with a 0.5ml syringe and the rest was removed leaving the pellet intact (nuclear fraction). 1ml of Trizol was wadded to the cytosolic fraction, while the nuclear fraction was first resuspended in 450µl of RNA Lysis Buffer 1X and 1mL of Trizol was subsequently added. RNA was extracted fallowing the Trizol manufacturer's instructions.

RNA-Fluorescence In Situ Hybridization (RNA-FISH)

For lncRNA *LUST* localization, RNA-FISH was performed fallowing Barakat's protocol with some modifications [164]. Cells were first transiently transfected with pcDNA3 or pcDNA3-*LUST* construct. 48 h post-transfections, cells were harvested and seeded on glass coverslips the day before the experiment. Cells were washed shortly with PBS and then fixed in 4% formaldehyde/5% acetic acid in PBS for 15 min. Subsequently, two washes with PBS and treatment with pepsin (0.1% in 10mM HCl) for 5 min. at 37°C, fallowed by two additional washes with water were performed. Cells were dehydrated using 70% ethanol for 3 min. 90% ethanol for 3 min., 100% ethanol for 3 min. and then let air dry for several minutes. 4.7µl of double-DIG labeled custom LNA mRNA detection probe (Stock concentration 84.4µM,

Exiqon) were put on a slide and covered with a 18x18 mm cover slip. The probe was denatured at 80°C for 75 sec, and the hybridization induced for 30 min. in a humid chamber, at hybridization temperature of 54°C (hybridization temp= Tm probe- 21°C). Cells were dehydrated as before and the hybridized probe was spot on a glass side and incubated with a coverslip on the top. Slides were placed in a humidified chamber filled with 50ml 50% formamide/2x Saline Sodium Citrate buffer (SSC: 3M NaCl, 300mM trisodium citrate pH 7), and incubated overnight at 37°C. The day after, a 6-well plate was filled with 2x SSC and pre-warm till 42°C. Coverlisps were added after overnight incubation and subsequently incubated at 42°C for 5 min. 2x SSC was removed and a further incubation with 50% formamide/2x SSC for 10 min at 42°C was performed, for 3 times. Slides were washed with Tris-saline-Tween (TST) for 2 x 5 min. at room temperature. Coverlisps were incubated upside down for blocking with 50µl of Tris-Saline-BSA (TSBSA), during 30 min. at room temperature in a TST-humidified dark chamber, and then transferred to a 6-well plate for two washes of 5 min. each with TST. Incubation with 50µl of mouse anti-DIG antibody (1:100 in TSBSA, Abcam) first, and with anti-mouse secondary antibody then, was subsequently performed for 30 min. at room temperature in the same way. Cells were dehydrated through 70% ethanol for 3 min. 90% ethanol for 3 min., 100% ethanol for 3 min. and let air dry for several minutes. Finally, a drop of mounting media mixed with DAPI, was used to mount the coverslips on each glass slide, sealed with nail polish and analyzed with Leica DM6000 microscope.

RNA Immuno-Precipitation (RIP) assay

Briefly, RNA is first cross-linked to proteins by formaldehyde, protein of interest is immune-precipitated by a specific antibody, then the cross-linking is reversed, and RNA is isolated and prepared for analysis [165]. In particular, Dynabeads® Protein G for Immuno-precipitation (Invitrogen), were washed twice and resuspended in 500µl RIPA buffer (150 mM NaCl, 1.0% NP-40, 0.5% Sodium deoxycholate, 0.1% SDS, 50mM Tris–HCl (pH 7.4), 1.0mM EDTA). 5µg of antibody against β-Catenin (Rabbit, Bethyl Laboratories) and non-specific IgG (Rabbit, Abcam) were added to the beads and incubated for 2 h at 4 °C with gentle rotation. HT-29 and CaCo2 cells were grown at ~90% confluence and harvested by trypsinization. An equivalent amount of media was added to quench the reaction and cells were collected in a 15ml conical tube and pellet by centrifugation at 500 x *g* for 10 min. Cells were washed twice with PBS and collected after each wash by pelleting at 500 x *g* for 5 min. The pellet was resuspended in 10 ml PBS and cells were counted using Trypan Blue. Cross-linking was performed by adding 37% formaldehyde to a final concentration of 0.3%, incubating the plates for 10 min., with gentle rotation at room temperature. The cross-linking reaction was quenched by adding 1.25M glycine to a final concentration of 0.125M, incubating the plates at room temperature for 5 min. Cells were collected and pelleted by centrifugation at 500 x *g* for 5 min, then washed twice, each time with 10ml of 1X PBS, spin and pellet as before. The pellet was resuspended in 2.2ml of RIPA buffer, supplemented with protease (Complete EDTA-free Protease Inhibitor Cocktail tablets, Roche) and RNAse (Recombinant RNasin Ribonucelase Inhibitor, Promega) inhibitors. The lysis was performed incubating at 37°C for 30 min., vortexing every 5 min. for 30 sec. intervals for the duration of the incubation. The sample was homogenized using a dounce homogenizer and the lysate was centrifuged at maximum speed (≥10,000 x *g*) using a microcentrifuge for 10 min. Supernatant was collected and the pellet

discarded. Lysate was pre-cleared using 30µl of protein G beads for 1h with gentle rotation at 4°C. Beads previously incubated with antibody were washed twice with 500µl of RIPA buffer and a 100µl aliquot of lysate was set apart to be used as Input sample. The rest of the lysate was divided between the two antibody-mounted beads (1ml each) and incubated overnight at 4°C with gentle rotation. The day after, the supernatant was removed by using a magnetic rack and the beads were washed four times with high salt RIPA buffer (1.0 M NaCl, 1.0% NP-40, 0.5% Sodium deoxycholate, 0.1% SDS, 50mM Tris–HCl (pH 7.4), 1.0mM EDTA). 50µl of sample were set apart to ensure that the antibody used worked and western blot analysis was performed. The remaining beads were resuspended in 100µl of Buffer C (150mM NaCl, 50mM Tris–HCl (pH 7.4), 5mM EDTA, 10mM DTT, 1.0% SDS), supplemented with 10µg of proteinase K (Fisher Bioreagents); 100µl of buffer C and proteinase K were also added to the input sample, and the reactions were incubated for 30 min. at 42°C for proteinase K digestion. A subsequently incubation for 4h at 65°C was performed to reverse the formaldehyde cross-links. After reverse cross-linking, 400µl of Trizol were added to the samples and RNA extraction was performed according to the manufacturer's protocol. cDNA synthesis was then performed as previously described.

UV Cross-Linking and Immuno-Precipitation (CLIP) assay

CaCo2 cells were cultured under normal growth conditions. Following UV cross-linking isolation of chromatin was performed using standard nuclear chromatin isolation techniques [166]. CLIP was performed essentially as described previously [167] with modifications described below. CaCo2 cells were harvested by trypsinization and resuspended in 1ml PBS, 1ml of nuclear suspension buffer (1.24M sucrose; 40mM Tris-HCl pH 7.5; 20mM $MgCl_2$; 4% Triton X-100), and 3ml water on ice for 20 min. Nuclei were pelleted by centrifugation at 2,500 x g for 15 min. and resuspended in a CLIP buffer (150mM KCl, 25mM Tris pH 7.4, 5mM EDTA, 0.5mM DTT, 0.5% NP40, 9μg/ml leupeptin, 9μg/ml pepstatin, 10μg/ml chymostatin, 3μg/ml aprotinin, 1mM PMSF, 100U/ml RNAse inhibitor; Ambion). The nuclear suspension was sheared using a bounce homogenizer with the B pestle using 10 strokes followed by 5 minutes of sonication (Bioruptor, Diagenode) at 4°C. Nuclear debris was pelleted by centrifugation at 15,000 rpm for 10 min. Once DNA tags were obtained and found within the appropriate nucleotide size range, samples were subjected to RNase-free DNA hydrolysis (Turbo DNAse, Ambion) and protein-RNA complexes were subjected to immuno-precipitation for CLIP. 100mg of antibodies for Cbx7 (Barradas et al., 2009), EZH2, β-Catenin, (Bethyl Laboratories), or TCF4 (Bethyl Laboratories) were added separately to supernatant and incubated for 2 h at 4°C with gentle rotation. Forty microliters of protein A/G beads were added and incubated for 1 h at 4°C with gentle rotation. Beads were pelleted at 2,500 rpm for 30 sec., the supernatant was removed, and beads were resuspended in 500 ml RIP buffer and repeated for a total of three washes in CLIP buffer, followed by one wash in PBS. Chromatin-bound RNA was then extracted from protein using Trizol and subsequently used for cDNA synthesis.

Statistical Analysis

All experiments were performed in triplicate at least 3 times. All values were expressed as mean ± SEM. Statistical analysis was performed by the Unpaired Student's *t* test. A probability value of $p < 0.05$ was considered statistically significant.

CHAPTER III: RESULTS

3.1 CD24 and CD44 expression defines the CCICs subpopulation in HT29 colon cancer cell line, and is upregulated during colonospheres formation

Previous studies showed that the detection of surface antigens such as CD24, CD44, CD133 and CD166 can be used to characterize and isolate the CCICs population in colon cancer cell lines [21,23,33]. Therefore, I first confirmed the expression of these stemness markers on 3 different colon cancer cell lines chosen for their differentiated status: SW480 (poorly differentiated), HT29 (moderately differentiated) [168] and LS174T (well differentiated) [169]. As shown in **Figure 1.3**, within the HT-29 cell line is it possible to identify a single subpopulation co-expressing all four stemness markers with bright intensity (CD24, CD44, CD133 and CD166, **Fig. 1.3**, *top right panel*, green subpopulation previous gated on $CD24^{Bright}/CD44^{Bright}$ intensity, *top left panel*). The same gating strategy was applied to LS174T and SW480 colon cancer cell lines, and as it shown, two different subpopulations could be discriminated based on the intensity of CD24 and CD44 co-expression (*middle and bottom left panels*), but the intensity of CD133 and CD166 antigens results to be lower (**Fig. 1.3** *middle and right panels*) compared to that of the HT-29 subpopulation. Taken together these results show that HT-29 cell lines could represent a better tool to study the role of the cancer initiating cells in colon cancer.

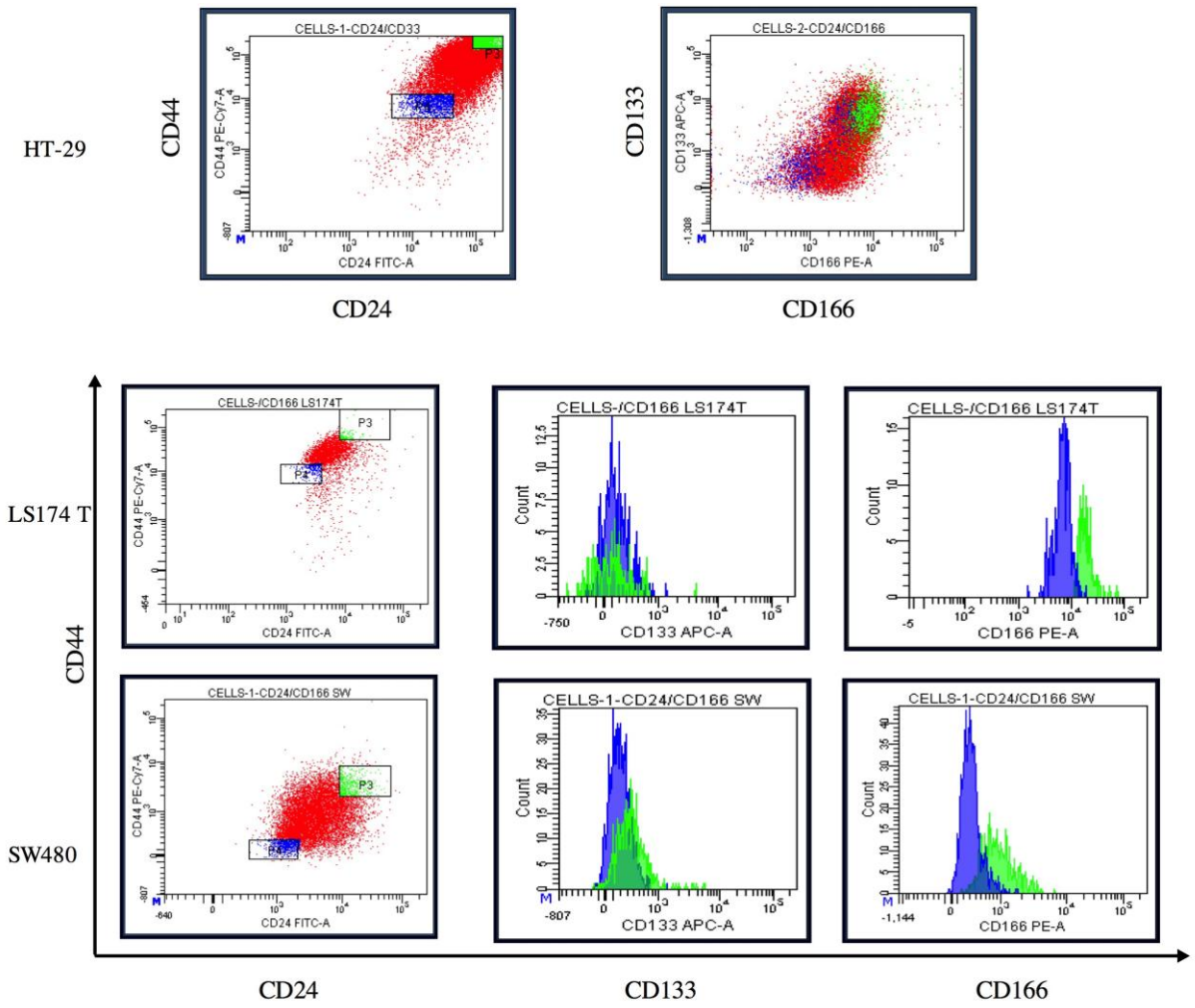

Figure 1.3 : Fluorescence-activated cell sorting (FACS) analysis

Intensity expression of CD24 (FITC), CD44 (PE-Cy7), CD133 (APC) and CD166 (PE) surface antigens on 3 different adherent colon cancer cell lines: HT29, LS174T and SW480. Histograms report the mean fluorescence intensity (MFI, range 0-10,000) of CD133 and CD166.

To assess the ability of self-renewal of differentiated *versus* undifferentiated subpopulations in HT-29 cells, I performed *in vitro* colonospheres formation assay as previously described [1,33]. To this purpose, I isolated HT29 $CD24^{bright}/CD44^{bright}$ (CCICs) and $CD24^{dim}/CD44^{dim}$ (the more differentiated counterpart) by Fluorescence-activated cell sorting (FACS) (**Fig. 2.3**).

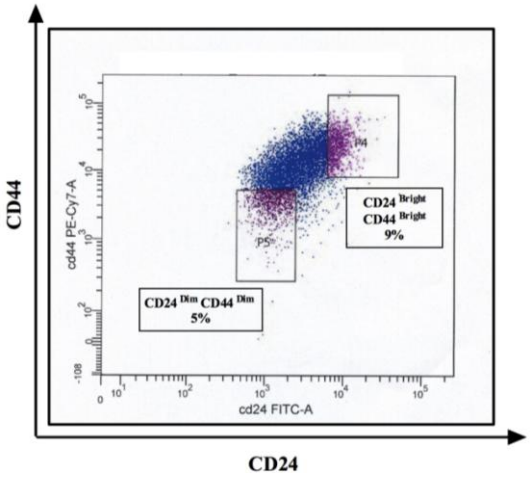

Figure 2.3 : Sorting of CCICs and more differentiated counterpart in HT-29 cancer cell line

Populations were gated using the CD24 (FITC) and CD44 (PE-Cy7) Bright and Dim expression. The percentage of selected subpopulations is shown.

Sorted cells were then growth in suspension, in serum-free media, to induce colonospheres formation [1]. **Figure 3.3** shows that the $CD24^{bright}/CD44^{bright}$ subpopulation is able to form more colonospheres compared to the $CD24^{dim}/CD44^{dim}$ subpopulation. Of note, colonospheres obtained from CCICs are larger in terms of number and size, compared to the ones obtained from the more differentiated counterpart (**Fig. 3.3**, *left and right panel, respectively*).

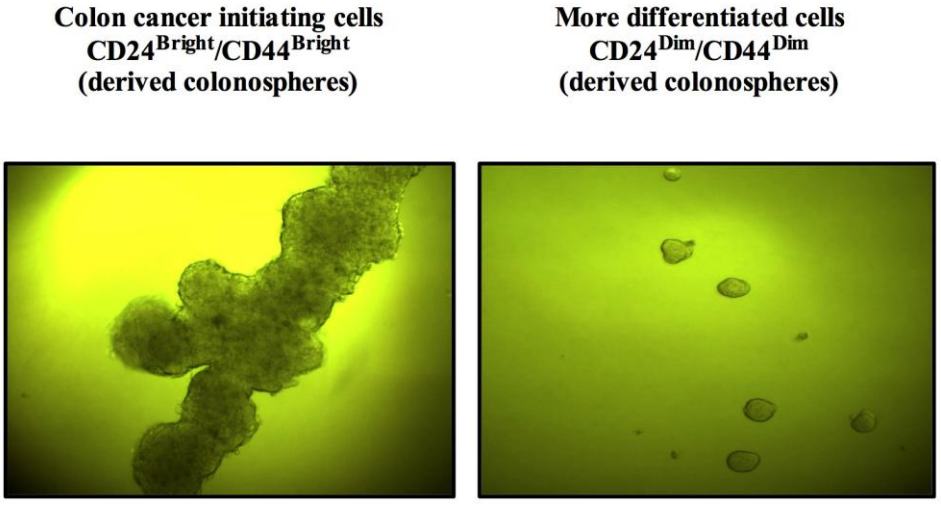

Figure 3.3 : HT-29 Colonospheres obtained from CCICs and the more differentiated counterpart

Left: colonospheres obtained from CCICs; *right:* colonospheres obtained from the more differentiated counterpart. Cells are cultured in ultra-low adherent conditions.

Taking together, these observations confirmed that the bright expression of CD24 and CD44 stemness markers could be used to isolate colon cancer initiating cells from HT-29 colon cancer cell line [33].

3.2 *Wnt*-target genes are induced within CCICs compared to the more differentiated counterpart

It is largely accepted that the *Wnt*/β-catenin signaling is the most deregulated pathway in colon cancer. Furthermore, previous studies have shown that this signaling regulates, specifically, the growth and maintenance of colonospheres [170,171]; in particular, formation of colonospheres from colon cancer stem cells subpopulations is associated to an elevated *Wnt*-signaling activity [171]. Based on these findings, I performed RNA-Seq on $CD24^{bright}/CD44^{bright}$ (CCICs) and $CD24^{dim}/CD44^{dim}$ (the more differentiated counterpart) subpopulations from HT-29 colon cancer cell line, and on colonospheres derived from the $CD24^{bright}/CD44^{bright}$ subpopulation, to explore pathways that could be eventually deregulated. Results showed that approximately 3000 genes were elevated in CCICs compared to the more differentiated counterpart, indicated with a fold change >1.5 (**Fig. 4.3**, *left upper panel*). The Gene Ontology analysis of the elevated genes revealed that a majority of them are involved in regulation of cell proliferation, response to hormone stimulus (hormone receptors), regulation of programmed cell death, positive regulation of cell migration and other biological processes characteristic of cancer cell phenotype (**Fig. 4.3**, *left lower panel*). The results also displayed that approximately 2000 genes are down-regulated between $CD24^{bright}/CD44^{bright}$ and $CD24^{dim}/CD44^{dim}$ subpopulations (**Fig. 4.3**, *upper right panel*), and gene ontology analysis revealed that a majority of them are also involved in regulation of cell proliferation, regulation of system process or other mechanisms involved in metabolism (**Fig. 4.3**, *lower right panel*).

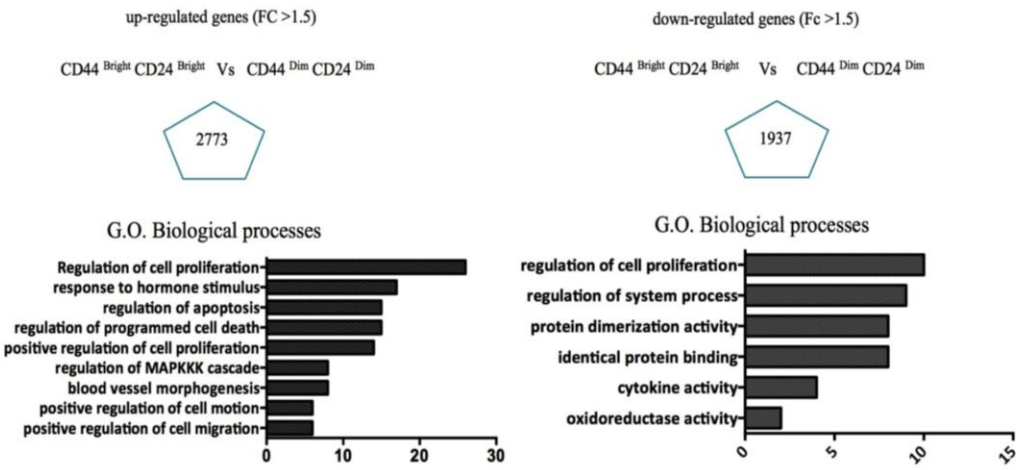

Figure 4.3: RNA-Seq analysis of CD24^{Bright}/CD44^{Bright} and CD24^{Dim}/CD44^{Dim} subpopulations

Upper panels: number of induced or up-regulate (*left*) and down-regulated (*right*) genes with a >1.5 fold change between CD24^{Bright}/CD44^{Bright} and CD24^{Dim}/CD44^{Dim} subpopulations. ***Lower panels***: gene ontology analysis of induced or up-regulated (*left*) and down-regulated genes (*right*).

To gain more detail of the characteristics linked with stemness *versus* differentiation, I analyzed the expression of genes defining differentiation or stem cell character, and the results showed that genes involved in mesenchymal cells differentiation, such as Mucin 2 (*MUC2*) [172] or Keratin B20 (*KB20/KRT80*) [173] are down-regulated, while colon cancer stem cells markers, in particular *CD24, CD44, CD166* and *ALDH1A1* are up-regulated (**Fig. 5.3**, *shown as the green and red bars, respectively*). Furthermore, members of the canonical *Wnt*-mediated signaling such as *ASCL2, IGFBP2, LGR4, DKK1, MYCL1, FGFR2, SP5, MMP7* and the receptor *EPHB3* were elevated in the CCICs compared to the more differentiated counterpart, with a fold increase >1.5 (**Fig. 5.3**, *shown as the blue bars*).

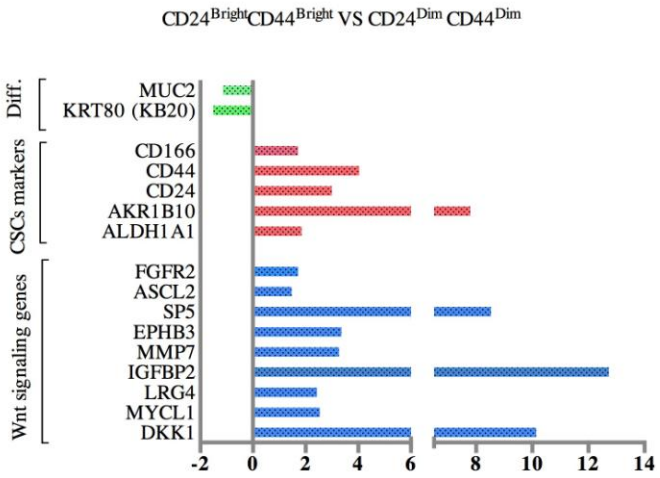

Figure 5.3: Expression of differentiation genes, CSCs markers and *Wnt*-signaling genes

Differential expression of differentiation genes (green), CSCs markers (Red) and *Wnt*-signaling genes (blue) between $CD24^{Bright}/CD44^{Bright}$ and $CD24^{Dim}/CD44^{Dim}$ subpopulations. Fold change >1.5 *p<0.01*

Next, I validated and confirmed through qRT-PCR, the overexpression of some of the most relevant *Wnt*-mediated target genes that play a pivotal role in colon cancer stem cells maintenance, and in particular the ones that regulate proliferation as Cyclin D1 (*CCND1*) and *MYC*, and the leucine-rich repeat containing G protein-coupled receptor 4 (*LGR4*), strong *Wnt*-mediated enhancer (**Fig. 6.3**).

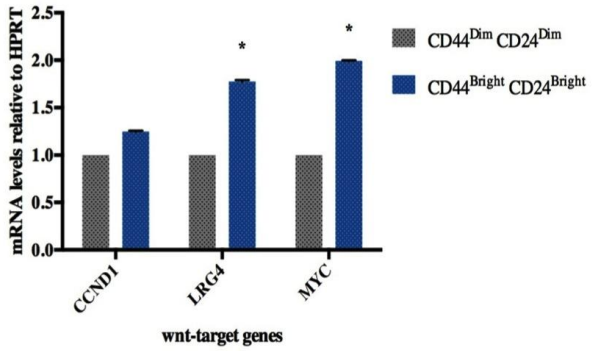

Figure 6.3: *Wnt*-mediated target genes expression between CCICs and differentiated cells

qRT-PCR validation of *Wnt*-mediated target genes differential expression between CCICs and the more differentiated counterpart. HPRT was used as housekeeping control gene for normalization. Error bars indicate the standard error of the mean (SEM). *$p<0.05$*

I also performed RNA-Seq on colonospheres and compared the expression of *Wnt*-target genes and stemness markers with the parental subpopulation ($CD24^{bright}/CD44^{bright}$) they were originated from. My results showed that the expression of *ASCL2, MYCL1, LGR4, EPHB3, CLDN1, CD44, CD24, ALDH1A1* and *AKR1B10* was strongly increased during colonospheres formation. What became of interest to me was that genes involved in differentiation, i.e. Krueppel-like factor 9 (*KLF9*) [174], *KB20* [173] and the β-catenin regulated gene Peptidyl Arginine Deiminase, Type 1 (*PADI1*) [175,176], were the most down-regulated (Fold change >1.5) gene transcripts in both, colonospheres and the parental subpopulation, as well as the retinoic acid receptor responder protein 1 (*RARRES1*), known to have a tumor suppressor role [177]. Intriguingly, two lncRNAs: *Kcnq1ot1* and *LUST*, resulted to be increased in this process, although the overexpression of *LUST* seemed to be more interesting because specific of the colonospheres formation process (**Fig. 7.3**).

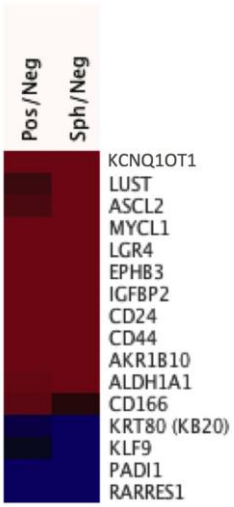

Figure 7.3: Heat map showing upregulated and down-regulated genes in CCICs

A Heat map showing up-regulated or induced genes (red) and down-regulated genes (blue) in $CD24^{Bright}/CD44^{Bright}$ compared to $CD24^{Dim}/CD44^{Dim}$ subpopulations, and in colonospheres compared to $CD24^{Dim}/CD44^{Dim}$ subpopulation. Fold change >1.5 $p<0.01$

These results show that *Wnt*-signaling is highly active during colonospheres formation. Therefore, I speculated that the expression of *Wnt*-target genes should be reduced in FBS-induced differentiation cells. To confirm this hypothesis, I evaluated the expression genes directly responsive to *Wnt*-signaling in colonospheres and FBS-induced differentiation. **Figure 8.3** shows that the mRNA levels of *CCDN1, LGR4* and *MYC* were reduced when cells were induced to differentiate (**Fig. 8.3**).

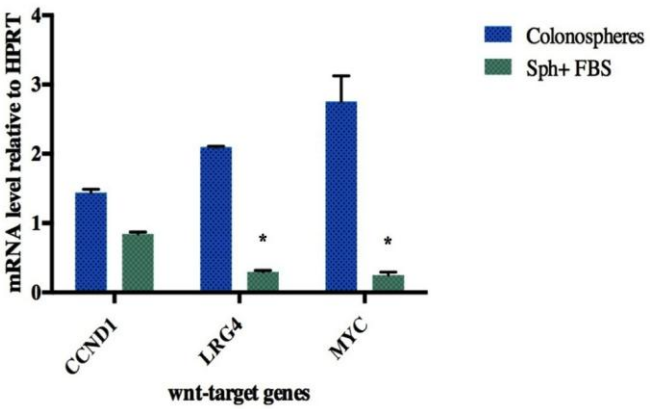

Figure 8.3: *Wnt*-mediated target genes expression in colonospheres differentiated cells

Validation by qRT-PCR of *Wnt*-mediated target genes differential expressed in colonospheres and FBS-induced differentiation cells. HPRT was used as housekeeping control gene for transcript normalization. Error bars indicate the standard error of the mean (SEM). *$p<0.05$*

Taking together, these observations show that the *Wnt*-signaling is strongly active during colonospheres formation and that the activity is reduced upon FBS-induced differentiation confirming its role in colonospheres maintenance [170].

3.3 The stemness capability of CCICs is strongly related to lncRNA *LUST* expression

Recent data have shown that deregulation of coding as well as non-coding RNAs, such as miRNAs and lncRNAs can contribute to cancer initiating cells generation (Reviewed in [178-180]). Lately, it has been demonstrated that, in cancer, the *Wnt*-signaling can be additionally regulated by lncRNAs through cell-autonomous mechanisms [181-183]. In order to investigate the involvement of lncRNAs in colon cancer initiating cells maintenance and their role in the *Wnt*-signaling regulation, I first performed a lncRNAs profiler array analysis on CCICs and the more differentiated counterpart isolated from HT-29 colon cancer cell line. Results from the array show that several lncRNAs were upregulated in CCICs compared to the differentiated counterpart**. Table 1.3** shows the most upregulated lncRNAs in CCICs, with a fold increase higher than 2.5.

Name	F.C.	Function
IGF2AS (family)	3.4	Unknown
Malat1	3.8	Metastasis-associated lung adenocarcinoma transcript1
PRINS	4	Required for cell viability after serum starvation
KRASP1	4.5	Unknown
LUST	4.52	Antisense to the putative tumor suppressor gene RBM5
Hoxa11as	4.6	Unknown
NDM29	4.8	Neuroblastoma differentiation marker 9
Kcnq1ot1	5.4	Involved in epigenetic silencing of the Kcq1 imprinting control region
H19 upstream conserved 1 & 2	5.8	Mesodermal enhancer
AntiPeg11	5.8	Binds PRC2
Jpx	6.6	Regulation of X Chr. Inactivation in mammals via activation of the Xist RNA
Tsix	7.3	Interacts with PRC2 complex
lincRNA-p21	7.5	Repress many genes transcriptionally regulated by p53
7SL	8.1	Scaffold SRP proteins
NEAT1(family)	9.5	Assembly, maintenance and structural integrity of paraspeckles
7SK	12.6	Transcription regulation
HOXA6as	18.3	Unknown

Table 3.1: lncRNAs up-regulated in CCICs compared to the differentiated counterpart

Among these lncRNAs, *LUST* was found to be overexpressed also in colonospheres by RNA-Seq, suggesting a functional role played in the colonospheres formation (**Fig. 7.3**).

To address the functional role of *LUST* on cancer stem cell maintenance, I performed *in vitro* colonospheres formation assay and analyzed by qRT-PCR the lncRNA and stemness markers expression in colonospheres and FBS-induced differentiation cells. As shown in **Figure 9.3** (*left*), the lncRNA *LUST* is strongly increased, as expected from the RNA-Seq and lncRNA profiler array data. Of note, the increase appears to be significant already after 14 days in culture, when the stemness markers levels (**Fig. 9.3**, *middle and right*) are not increased yet, and reaches a 10 fold change increase when cells are grown for 5 weeks in the same conditions (**Fig. 9.3**, *left*). This result suggests that *LUST* upregulation could be an early event in the colonospheres formation.

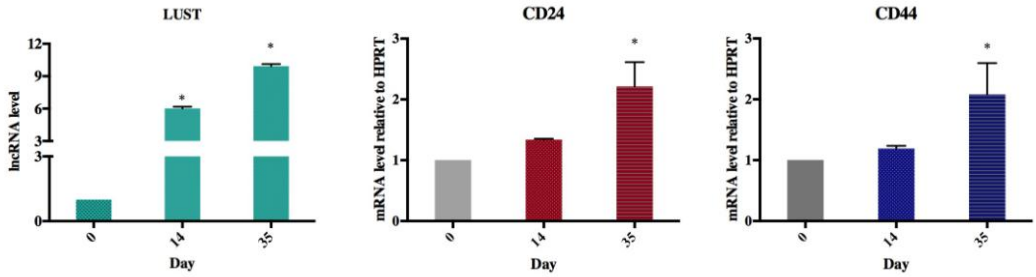

Figure 9.3: *LUST* lncRNA, CD24 and CD44 mRNA expression levels in HT29 colonospheres

LUST lncRNA, CD24 and CD44 mRNA levels in HT29 colonospheres, respectively after 14 and 35 days cultured in ultra-low attachment conditions. HPRT was used as housekeeping control gene. Error bars indicate the standard error of the mean (SEM). *$p<0.05$.

As expected, when colonospheres differentiation is induced by adding FBS [23], the mRNA expression levels of stemness markers CD24 and CD44 is abrogated (**Fig. 10.3**, *left and middle panel*), confirming the loss of the stemness potential. Interestingly, the expression level of *LUST* drastically decreases (**Fig. 10.3**, *right panel*).

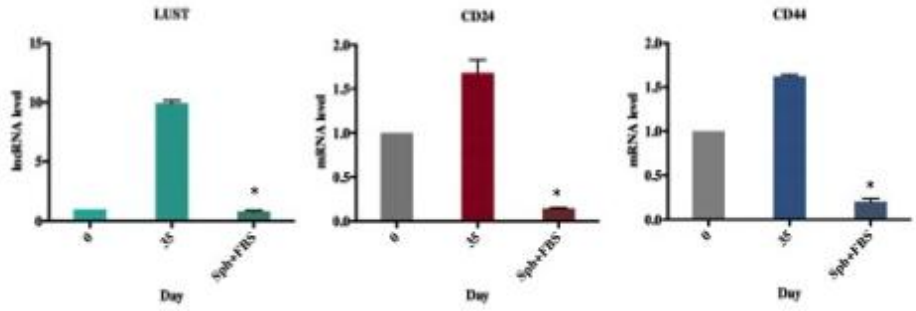

Figure 10.3: *LUST* lncRNA CD24, CD44 mRNA levels in FBS-induced differentiation cells

LUST lncRNA, CD24 and CD44 mRNA levels in colonospheres and in FBS-induced differentiation cells. Gene relative levels were analyzed by qRT-PCR. HPRT was used as housekeeping control gene for normalization. Error bars indicate the standard error of the mean (SEM). *$p<0.05$

Furthermore, the loss of stemness markers is also shown at the protein level (**Fig. 11.3**, *blue population*).

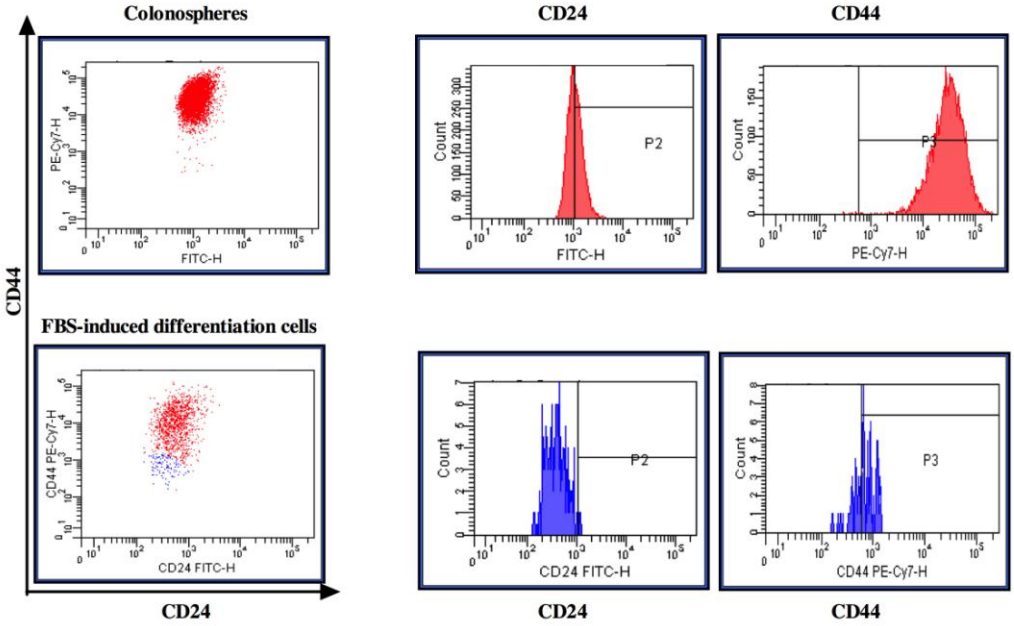

Figure 11.3: FACS analysis of Colonospheres and FBS-induced differentiation cells

FACS analysis of Colonospheres (*upper panels*) and FBS-induced differentiation cells (*lower panels*) showing loss of expression of stemness markers (blue population). Histograms report the mean fluorescence intensity (MFI, range 0-10,000) of CD24 (FITC) and CD44 (PE-Cy7).

3.4 *LUST* knock-down impairs *Wnt*-target genes activation

In order to investigate *LUST* involvement in *Wnt*-signaling, I silenced *LUST* expression in HT29 cells using locked nucleic acid (LNA) longRNA GapmeRs, which are single stranded antisense oligonucleotides (ASO) used for efficient inhibition of lncRNAs expression (Exiqon). **Figure 12.3** shows, respectively, the 60% and 50% loss of *LUST* expression using two different ASOs, measured by qRT-PCR (**Fig. 12.3 A**). Interestingly, LncRNA *LUST* knockdown also reduces cells growth as shown in **Figure 12.3 B**.

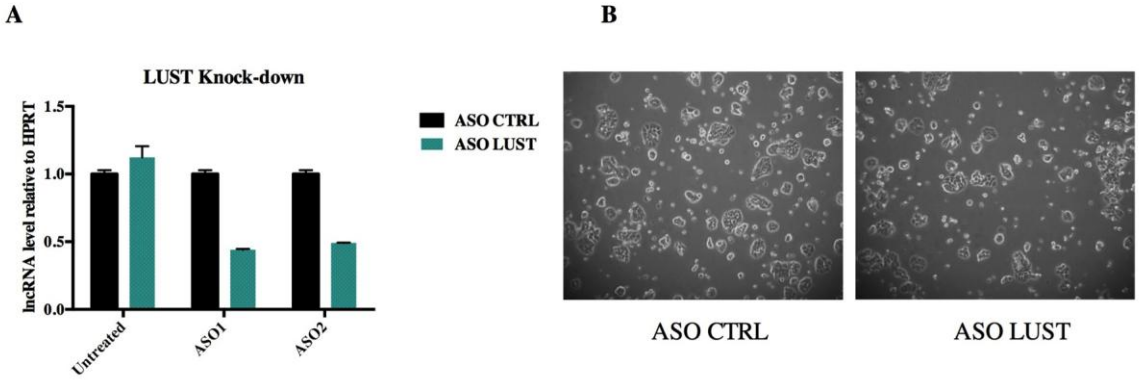

Figure 12.3: lncRNA *LUST* knockdown

A) Knock-down of the lncRNA *LUST*. lncRNA expression is reduced, respectively, to 40% and 50% from ASO1 and ASO2. An ASO unspecific probe was used as negative control (ASO CTRL). HPRT was used as housekeeping control gene for normalization. Error bars indicate the standard error of the mean (SEM). **B)** Cell growth reduction upon lncRNA *LUST* knockdown.

To verify an involvement of *LUST* in *Wnt*-signaling regulation I analyzed the expression of *Wnt*-signaling target genes. My results showed that, interestingly, a strong reduction of several target genes such as *AXIN2*, *CCND1*, *CD44*, and *TCF4* as well as *CD24* mRNA levels occurs upon *LUST* knockdown (**Fig. 13.3**).

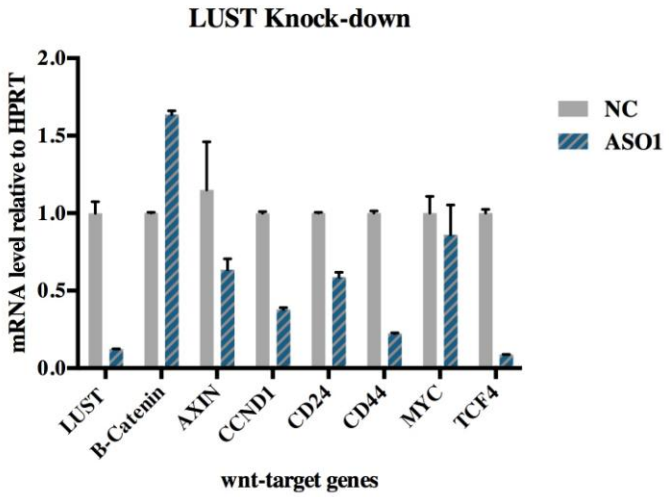

Figure 13.3: *Wnt*-signaling target genes expression in *LUST* knockdown cells

mRNA levels of *Wnt*-mediated genes upon *LUST* knockdown were analyzed by qRT-PCR. HPRT was used as housekeeping control gene for normalization. Error bars indicate the standard error of the mean (SEM).

Furthermore, **Figure 14.3** shows a decrease also at the protein level of CCND1 and c-Myc. Of note, *LUST* knockdown caused a reduction of active β-catenin (**Fig.14.3**).

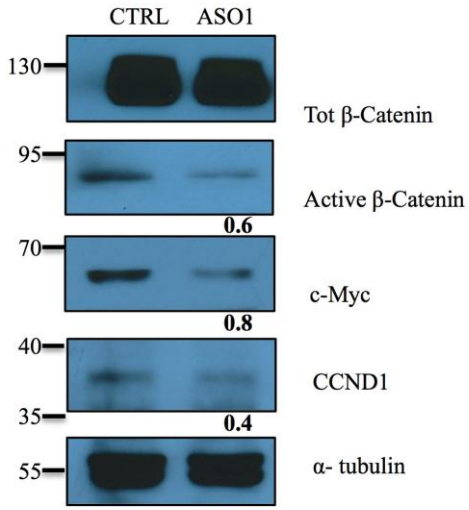

Figure 14.3: *Wnt*-signaling member proteins levels in *LUST* knockdown cells

WB analysis of total β-catenin, active β-catenin, c-Myc and CCND1 in ASO control cells and in *LUST* knockdown cells. 30μg of protein from total cell lysate were analyzed. α-tubulin was used as a loading control.

Taken together, these results demonstrate that *LUST* regulates *Wnt*/β-catenin signaling.

3.5 *LUST* overexpression induces *Wnt*/β-catenin signaling activation

I performed dual luciferase assay using M50 Super 8x TOPFlash promoter, which contains two sets of 3 copies of the wild-type TCF binding regions upstream of a luciferase reporter, transiently integrated in HT29 cells (TOPFlash-HT29 cells). Relative luciferase activity was measured in TOPFlash-HT29 cells transfected with pcDNA3-*LUST*, pcDNA3-β-catenin, or pcDNA3-*LUST* and pcDNA3-β-catenin.

The results showed that the lncRNA *LUST* alone as well as β-catenin is able to activate TCF reporter. Interestingly, the co-transfection of pcDNA3-*LUST* and pcDNA3-β-catenin induced a really strong increase in luciferase activity. These results demonstrated that the lncRNA *LUST* and β-catenin synergistically activate *Wnt*/β-catenin signaling (**Fig. 15.3**).

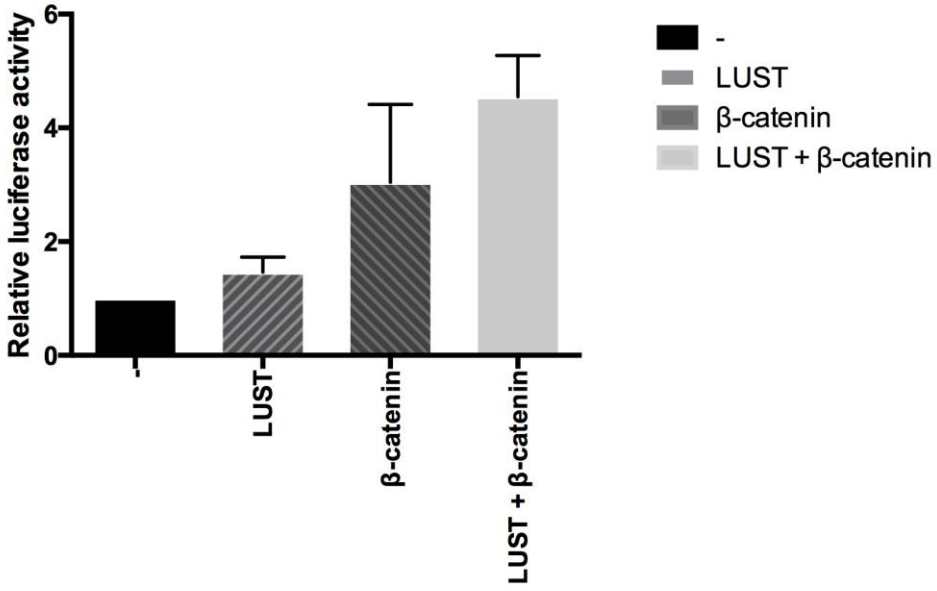

Figure 15.3: TOPFlash Luciferase Assay in *LUST* overexpressing cells

Relative luciferase activity measured in TOPFlash-HT29 cells transiently transfected with pcDNA3-*LUST*, pcDNA3-β-catenin, or pcDNA3-*LUST* and pcDNA3-β-catenin.

To further evaluate the *LUST*-mediated regulation of *Wnt*-signling in colon cancer, I analyzed the mRNA levels of *Wnt*-signaling target genes upon *LUST* overexpression. pcDNA3 and pcDNA3-*LUST* constructs were transiently transfected into HT29 cells and RNA samples were analyzed through qRT-PCR. The results show the expression of *AXIN2, CCND1* and *TCF4* increased with a fold change of 2.8. The mRNA levels of the stemness markers *CD24* and *CD44* are also upregulated in these conditions, with a fold increase of 2. Interestingly, the mRNA expression level of *MYC* is approximately 4 times higher compared with its expression in cell transfected with empty vector. These results confirm that *LUST* activates the *Wnt*-signaling in colon cancer cells (**Fig. 16.3**).

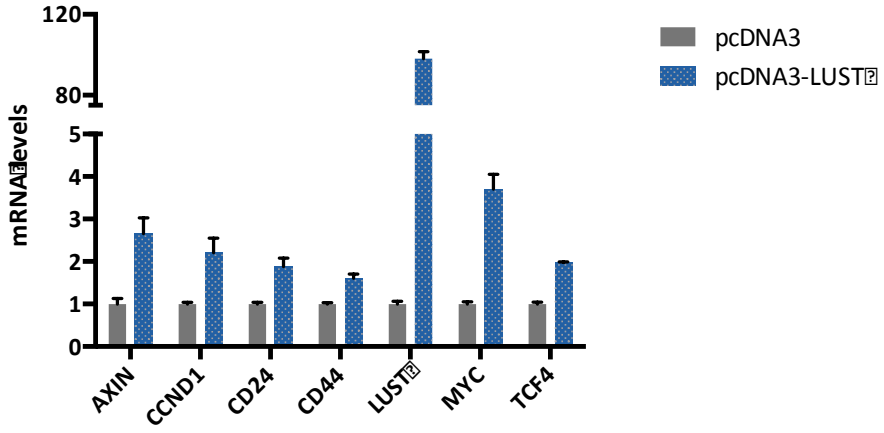

Figure 16.3: *Wnt*-signaling target genes expression in *LUST* overexpressing cells

mRNA levels of *Wnt*-target genes upon *LUST* overexpression were analyzed by qRT-PCR. HPRT was used as housekeeping control gene for normalization. Error bars indicate the standard error of the mean (SEM).

Once I demonstrated that *LUST* overexpression induces activation of *Wnt*-signaling target genes transcription, I further investigated if the increase at the transcriptional level was reflected by a corresponding increase at the protein level. To this purpose, I analyzed by western blot the hypothetical changes of protein levels in lncRNA *LUST* overexpressing cells. As shown in **Figure 17.3**, total levels of β-catenin remain stable, while there is an increase of expression of active β-catenin, MYC and CCND1, demonstrating that the increase in induced at both, mRNA and protein levels.

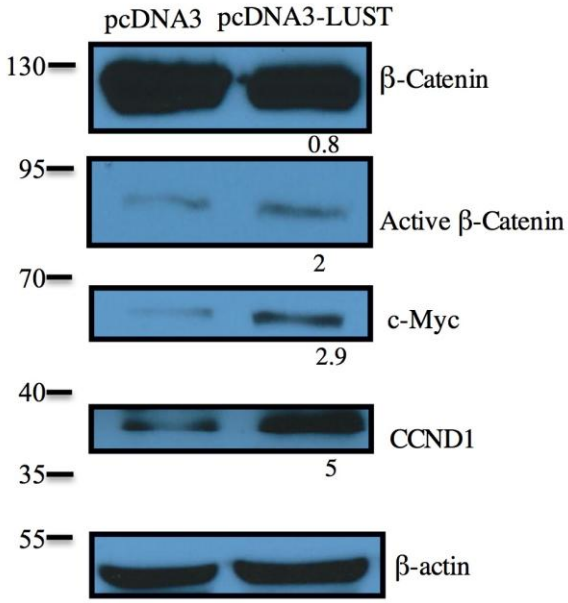

Figure 17.3: *Wnt*-signaling member proteins levels in *LUST* overexpressing cells

WB analysis of total β-catenin, active β-catenin, c-Myc and CCND1 in *LUST* overexpressing cells and in pcDNA3 control cells. 30μg of protein from total cell lysate were analyzed. β-actin was used as a lading control.

3.6 *LUST* is a nuclear lncRNA that strongly binds to β-catenin

To identify the molecular mechanism by which *LUST* regulates the *Wnt*-signaling, I performed nuclear/cytoplasmic RNA fractionation in HT29 colon cancer cell lines. My results indicated that *LUST* lncRNA localization is mostly nuclear in both cell lines, although a negligible amount of transcript is also detected in the cytoplasmic compartment of the cells (**Fig. 18.3**). Malat1, a nuclear lncRNA, was used as a positive control [184].

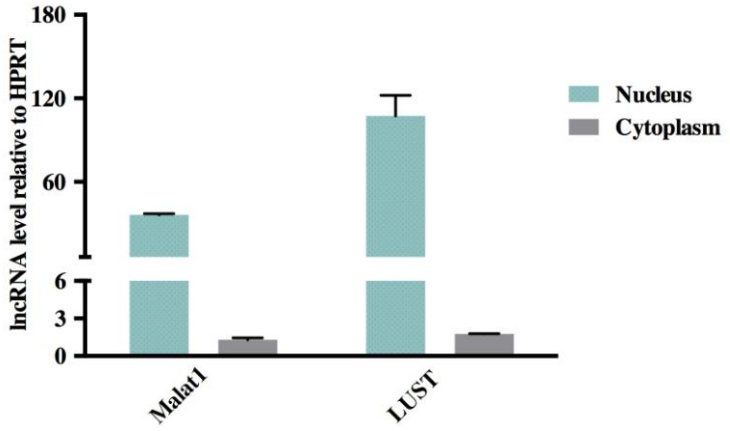

Figure 18.3: HT29 Nuclear/Cytoplasmic RNA fractionation

Bar plots show the relative abundance of *LUST* measured by qRT-PCR. *Malat1* was used as positive control for nuclear localization. HPRT was used as housekeeping control gene for normalization. Error bars indicate the standard error of the mean (SEM).

To further validate *LUST* cellular localization, we performed RNA Fluorescence In Situ Hybridization (RNA-FISH) in HT29 cells transiently transfected with pcDNA3 (empty vector) or pcDNA3-*LUST*. Results shown that in HT29 cells transiently transfected with pcDNA3-*LUST*, the transcript expression is mostly nuclear (red spots) (**Fig. 19.3**).

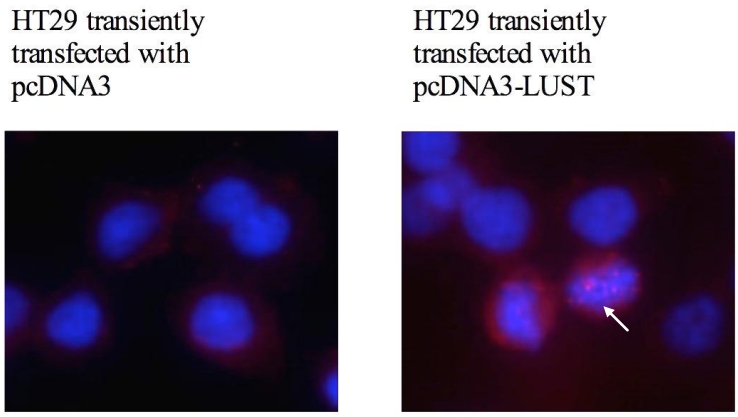

Figure 19.3: RNA-FISH in HT29 colon cancer cells

RNA-FISH on HT29 transfected with pcDNA3 (left) or pcDNA3-*LUST* (right). Nuclei are stained with DAPI (blue), lncRNA *LUST* was detected using LNA double-DIG mRNA probe (red). White arrow indicates the *LUST* transcript.

The involvement of lncRNAs in colorectal cancer has been largely demonstrated and in particular, several studies demonstrated the interaction of lncRNAs with *Wnt*/β-catenin signaling in cancer [185,186]. To analyze the role played by *LUST* in the regulation of β-catenin activity, I performed RNA Immuno-Precipitation (RIP) [165] in HT29 cell lines, immuno-precipitating β-catenin and using IgG as negative control. Results demonstrated that the lncRNA *LUST* strongly binds to β-catenin as shown in **Figure 20.3**.

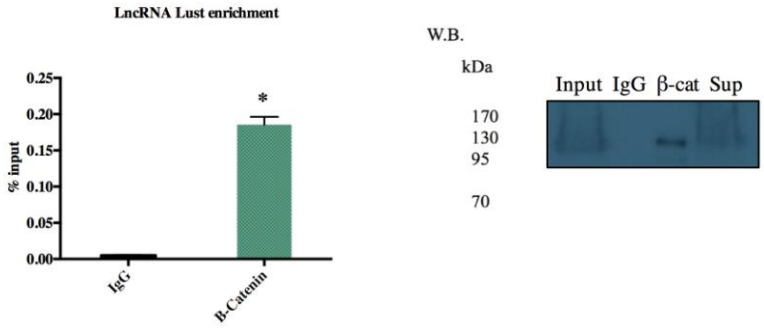

Figure 20.3: RNA Immuno-Precipitation (RIP) of β-catenin in HT29 cells

Left: Enrichment of lncRNA *LUST* binding to β-catenin protein. Enrichment is shown as % input. *Right:* RIP-WB to confirm β-catenin immuno-precipitation. IgG was used as negative control. Sup: supernatant. Error bars indicate the standard error of the mean (SEM). *$p<0.05$*

To further confirm the binding of *LUST* with β-catenin, I performed Cross-Linking Immuno-Precipitation (CLIP) assay [167] in CaCo2 colon cancer cell line, which harbors both, APC and β-catenin mutations [187].

The CLIP experiment was performed by immune-precipitation of β-catenin, the transcription factor 4 (TCF4), Histone-lysine N-methyltransferase EZH2, and the chromobox protein homolog 7 (CBX7). β-catenin and TCF4 were analyzed because members of *Wnt*-signaling [188], while EZH2 was used as positive control for lncRNA-binding [189] and IgG and CBX7 as negative controls for *Wnt*-signaling. Interestingly, we were able to detect the binding of lncRNA *LUST* with β-catenin and TCF4 with a 2.5 and 1.8 fold change, respectively, compared to the IgG or CBX7 binding, used as negative controls (**Fig. 21.3**).

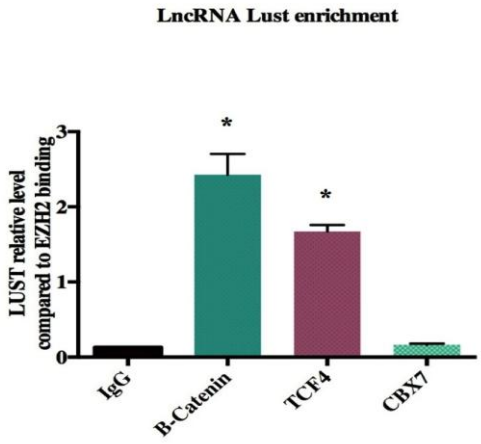

Figure 21.3: Cross-Linked Immuno-Precipitation (CLIP) assay in CaCo2 colon cancer cells

Enrichment of lncRNA *LUST* to β-catenin and TCF4 is shown. IgG and CBX7 were used as negative controls. Error bars indicate the standard error of the mean (SEM). *$p<0.05$

A supplemental validation of my results was obtained by performing RIP in CaCo2 as shown in **Figure 22.3**.

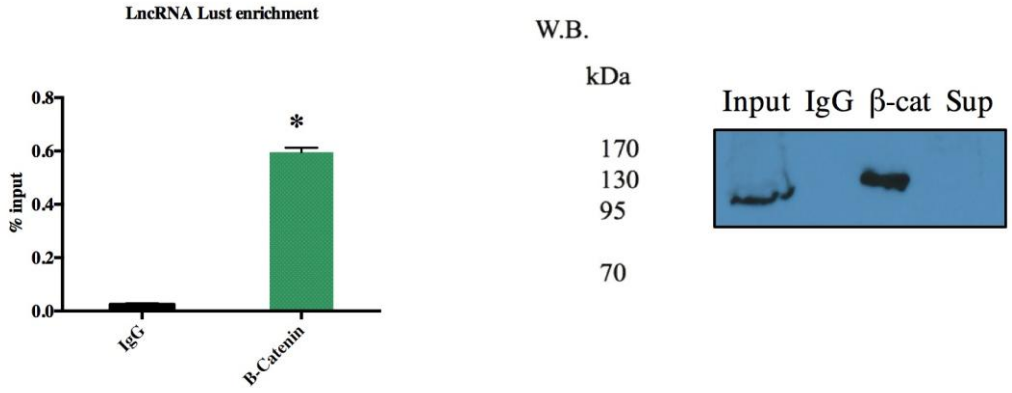

Figure 22.3: RNA Immuno-Precipitation (RIP) of β-catenin in CaCo2

Left: Enrichment of lncRNA *LUST* binding to β-catenin shown as % input. *Righ:* RIP-WB to confirm β-catenin immuno-precipitation. IgG was used as negative control. Sup: supernatant. Error bars indicate the standard error of the mean (SEM). *$p<0.05$*

3.7 *LUST* overexpression accelerates colonospheres formation of CCICs

To demonstrate that *LUST* could be the key effector in colonospheres formation, I performed *in vitro* colonospheres formation from HT29 cells upon *LUST* overexpression. **Figure 23.3** shows that *LUST* overexpression promotes an earlier colonospheres formation, compared to the control. In fact, HT-29 cells begin to form colonospheres after 21 days cultured in serum-free media. LncRNA *LUST* overexpressing cells, instead, are able to form colonospheres already after 7 days in the same culture conditions. The full colonospheres morphology is reached after only 14 days, instead of 21 days needed for the untransfected cells.

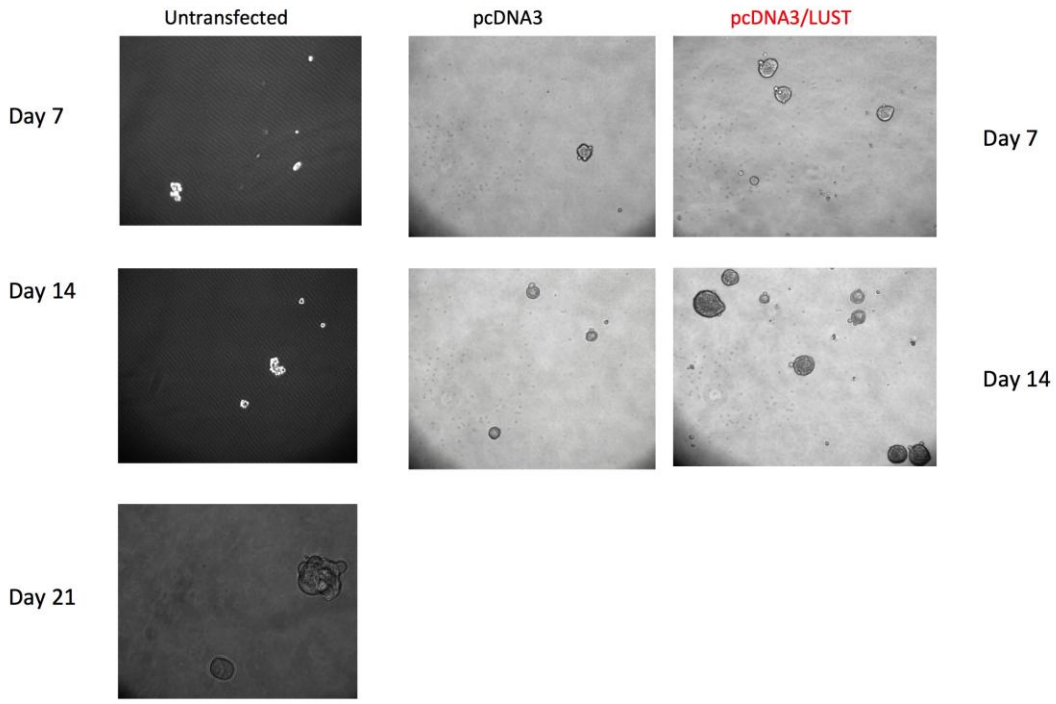

Figure 23.3: Colonospheres formation in *LUST* overexpressing cells

HT29 untransfected cells were cultured in ultra-low attachment conditions, in serum free media for 7, 14 and 21 days. pcDNA3, and pcDNA3-*LUST* transfected cells were cultured in the same conditions for 7 and 14 days. Colonopsheres are shown.

These results, strongly confirmed that *LUST* promotes colon cancer initiating cells maintenance, enhancing the activation of *Wnt*-signaling through β-catenin binding.

CHAPTER IV: DISCUSSION

Colorectal cancer (CRC) is the third most common cancer and the third cause of cancer death in the United States. This type of malignancy is considered one of the most complete examples of a hierarchically organized solid cancer dominated by a subpopulation of immature cells with specific molecular and functional features. Human colorectal cancer initiating cells (CCICs) were first isolated on the basis of CD133 expression and demonstrated to induce tumors in mice that recapitulate the original malignancy [22,23]. During the past years, several groups have expanded the search for other surface markers of CCICs in hopes of finding specific biomarkers, which could serve as prognostic and therapeutic targets. Several CCICs markers and combined phenotypes have been described, including CD44 [21,30,31], CD24 and CD29 [27] as markers, and $CD24^{+}CD44^{+}$ [33], $EphB2^{high}$ [190], $EpCAM^{high}/CD44^{+}/CD166^{+}$ [21], $ALDH^{+}$ [191], $LGR5^{+}$ [192], and $CD44v6^{+}$ [193] as phenotypes. In this study, I used the previously described and characterized "stem-like" markers, CD24, CD44, CD133 and CD166 to identify CCICs populations from three different colorectal cancer lines: SW480 (poorly differentiated), HT-29 (moderately differentiated) [168] and LS174T (well differentiated) [169] (**Fig.1.3**). FACS analysis showed that only within the HT-29 colon cancer cell line is it possible to identify a subpopulation characterized by the bright co-expression of all the four stemness markers evaluated, that likely represent the CCICs population. LS174T and SW480 colon cancer cell lines, do not show a bright intensity of CD133 and CD166 as HT-29 colon cancer cell line. This results can be explained by the differentiation potential of the cells, in fact low intensity of "stem-like" antigens expression on LS174T was expected considering that is a well differentiated colon cancer cell line [169]. Although SW480 is a poor differentiated colon cancer cell line [168], the low expression of these markers can be explained considering that not all the cells with a low differentiation state are cancer stem cells [7,23]; while a small population as the

one identified in HT-29 colon cancer cell line, with stemness potential and with bright intensity of CD24, CD44, CD133 and CD166 "stem-like" markers, can give rise to tumor initiation and formation [6,8,21]. In particular, I focused my attention on the HT-29 subpopulations expressing the $CD24^{Bright}/CD44^{Bright}$ and $CD24^{Dim}/CD44^{Dim}$ phenotype (**Fig.2.3**) that would represent, respectively, the CCICs, subpopulation and the more differentiated counterpart, as previously described by Yeung et al. [33]. However, several questions remain open in the field of colorectal CSC identification. In fact, the consistency of CSC-associated markers needs to be further investigated, given that the CSC phenotype itself has been shown to be unstable. For example, CCICs enriched populations that are positive and negative for LGR5 can interconvert upon chemotherapy [194]. Based on these findings, I further validated the stemness potential of the $CD24^{Bright}/CD44^{Bright}$ and $CD24^{Dim}/CD44^{Dim}$ isolated subpopulations. Neoplastic stem cells from neural and epithelial organs can be expanded as sphere-like cellular aggregates in serum-free medium containing EGF and FGF. Thus, colon undifferentiated cancer cells can be cultured and expanded *in vitro* as colon spheres in proliferative serum-free medium containing growth factors. This property is common to neural and epithelial stem and progenitor cells, which grow as spherical aggregates that in the presence of serum or extracellular matrix differentiate upon growth factor removal [20,195-197]. To this end, I cultivated both the subpopulations of cells obtained upon fluorescence-activated cell sorting, in such a proliferative medium for undifferentiated cells. After 3 weeks of culture, I obtained colonospheres formed by aggregates of exponentially growing undifferentiated cells $CD24^{Bright}/CD44^{Bright}$ (CCICs) that were larger in terms of size and number compared to the ones obtained from the $CD24^{Dim}/CD44^{Dim}$ colon cancer cells (the more differentiated counterpart) (**Fig. 3.3**). This result can be explained considering that CD24 and CD44 are stem-like markers [19,198] therefore, cells that express these markers at higher intensity are supposed to be enriched in stemness properties, thus able to grow in such conditions. In this

contest, it has been shown that cytokines produced by tumor-associated cells can induce increased CSC self-renewal [171,199,200], therefore have a role in CSCs maintenance as external effectors. In fact, Todaro et al. demonstrated that the cytokines hepatocyte growth factor (HGF), osteopontin (OPN), and stromal-derived factor 1α (SDF-1), secreted from tumor associated cells, increase CD44v6 expression in CCICs by activating the *Wnt*/β-catenin signaling, which promotes migration and metastasis [193]. Thus, it is possible that the percentage of cells expressing CSC markers within a tumor, may vary depending on disease stage, considering that a more strict hierarchical organization has been proposed to be present in the early stages and more relaxed in advanced stages of the disease [201]; also, this percentage may vary depending on the timing and type of therapy, and on a series of microenvironmental and individual factors that are predictably difficult to define. Therefore, in a dynamic scenario where CSCs may change in quantity and phenotype during tumor progression, the expression of CSC markers should be seen as a relative and contextual parameter rather than a general property of the tumor. Besides phenotypic markers, another way to identify CSCs is through their individual molecular and/or functional features. Interestingly, from the molecular side, the hallmark of colorectal CSCs has been shown to be a hyperactivated β-catenin pathway, which translates into the ability to generate serial tumors in vivo [171]. Therefore, my first hypothesis was that the phenotypic differences between the CCICs and the more differentiated counterpart could be individually reflected at the molecular and functional level. To this purpose I performed RNA-Seq of CCICs ($CD24^{Bright}/CD44^{Bright}$), colonospheres and more differentiated counterpart ($CD24^{Dim}/CD44^{Dim}$), to explore pathways that could be eventually *dys*-regulated (**Fig. 4.3**). I analyzed the expression of genes defining differentiation or stem cell character in CCICs and in the more differentiated counterpart, and the results showed that genes involved in colonic mesenchymal cells differentiation, such as Mucin 2 (*MUC2*) [172] or

Keratin B20 (*KB20/KRT80*) [202] were down-regulated in CCICs demonstrating that this population is less differentiated compared to the $CD24^{Dim}/CD44^{Dim}$. The increased expression of colon cancer stem cells markers, in particular *CD24, CD44, CD166* and *ALDH1A1*, confirms the stemness potential of the cells. What was really interesting, was that members of the canonical *Wnt*-mediated signaling such as *ASCL2*, *IGFBP2*, *LGR4, DKK1*, *MYCL1*, *FGFR2, SP5, MMP7* and the receptor *EPHB3* were elevated in CCICs compared to the more differentiated counterpart (**Fig. 5.3 and Fig. 6.3**). The further overexpression of *Wnt*-signaling pathway genes in colonospheres confirmed by RNA-Seq (**Fig. 7.3)**, corroborates the findings from Vermeulen et al [171]. Moreover, the experiments that I performed on FBS-induced differentiation cells showed that the expression of *Wnt*-signaling target genes is drastically reduced upon differentiation, adding a further confirmation to the involvement of this signaling in colonospheres formation (**Fig. 8.3**). Self-renewal is a stem cell functional trait, which has been shown in some studies on colorectal cancer stem cells to be depend on some transcriptional regulators such as isopentenyl-diphosphate delta isomerase 1 (ID1) and -3 (ID3)[203], or BMI1 that has also been shown to be a key player of self-renewal in CCICs, as its inhibition results in stem cell loss and impairment of tumor growth [204]. Therefore, we may consider the idea of a new point of view where stemness arises from the continuous adaptation of cancer cell populations to microenvironmental signals. Increasing evidence suggests that, in both the normal and neoplastic intestine, stemness results from the incessant convergence of cell-intrinsic features (genetic mutations and epigenetic regulation), local signals (of a chemical, mechanical, and molecular nature), stochastic events, and population forces that continuously shape the stem cell pool [205]. In particular, the epigenetic regulation of transcription is a complicated and comprehensive event, that can involve different players, and recent evidences revealed that long-noncoding RNAs can impersonate this role. Several

studies have shown that these lncRNAs can work as transcriptional activators or repressor, function as transcriptional guide, or act as scaffold for chromatin modification complex (reviewed in [106]). To make a more direct point, it has been shown that the lncRNA *ANRIL* physically interacts with CBX7, member of PRC1 complex; this interaction induce formation of heterochromatin surrounding the INK4b-ARF-INK4a locus, leading to its repression. This mechanism facilitates the bypassing of senescence, endowing the prostate cancer cells stemness [122,123]. Also, previous reports showed the involvement of lncRNAs in regulating cell stemness maintenance [206,207]. Moreover, recent findings showed that lncRNAs play significant roles in regulating cellular development and differentiation, processes that are frequently deregulated in cancer [208]. The involvement of lncRNAs in colorectal cancer has been demonstrated from several groups, and in particular, Nissan et al. were the first showing that CRC-associated transcript 1 (CCAT1) is highly expressed in CRC but not in normal tissue [209], while Xiang et al. demonstrated that this lncRNA plays a role in MYC transcriptional regulation and promotes long-range chromatin looping [182]. Interestingly, recent findings showed that, in cancer, the *Wnt*-pathway can be additionally regulated by lncRNAs [181-183]. In this scenario, considering that HT-29 colon cancer cell line harbors an *APC* inactivating mutation (deletion at the carboxyl terminus at residue 1555) that gives rise to a truncated and inactive APC protein [210-212], inducing the cells to be insensitive to external stimulation by Wnt ligands, but sensitive to the constitutively activation of the *Wnt*-signaling within the cell [213], I conceived that new kinds of transcriptional regulator, not fully investigated yet, such as lncRNAs, could have a role during the colonospheres formation process, thus in CCICs maintenance by regulating the *Wnt*-signaling. For this reason, I performed a lncRNA profiler array analysis on CCICs and the more differentiated counterpart subpopulations from HT-29 cell line, evaluating the differential expression of 90 different lncRNAs. The expression of several lncRNAs resulted to be increased in CCICs compared to the differentiated counterpart,

and in particular: *HOXA6as, 7SK, NEAT1* (family), *7SL, lincRNA-p21, Tsix, Jpx, AntiPeg11, H19, Kcnq1ot1, NDM29, Hoxa11as, LUST, KRASP1, Prins, Malat1* and *IGF2AS* (family) (**Table 1.3**). According to the comprehensive lncRNA consortium database (lncRNAdb) (http://www.lncrnadb.org), most of the lncRNAs have been characterized to a degree at the functional level, except for *HOXA6as, IGF2AS* (family) and *KRASP1* which role is still unknown. Among the upregulated lncRNAs, *LUST* was found overexpressed also in colonospheres (**Fig. 7.3**), suggesting a functional role played by this lncRNA during the colonospheres formation process. My data showed that only two lncRNAs were found increased in both, $CD24^{bright}/CD44^{bright}$ subpopulation and colonospheres, in particular, *Kcnq1ot1 (KCNQ1 opposite transcript 1)* and *LUST (RBM5-AS1)*. *Kcnq1ot1* has been widely characterized and has been demonstrated that this lncRNA has a role in the organization of a tissue/lineage-specific nuclear domain involved in epigenetic silencing of the Kcnq1 imprinting control region. To further detail this concept, the expression profile, function and epigenetic alterations of this locus in colorectal cancer have been previously characterized [116,214]. The functional role of the lncRNAs *LUST* (Luca-15 specific transcript/RBM5-AS1) has just been hypothesized. *LUST* is an antisense transcript of ~1.3kb that initiates within intron 6 and terminates within intron 4 of the RBM5 gene. Sutherland et al. speculated that *LUST* binds and masks an unknown sense-strand regulatory sequence, common to two *RBM5* different isoforms, preventing the premature termination of both mRNAs. In absence of *LUST*, the expression of this truncated isoform triggers apoptosis, therefore, conferring to *LUST* an antiapoptotic role [149]. However, the molecular function of *LUST* remains still unclear and further investigations regarding its role and expression in colorectal cancer have never been pursued. The *Wnt*-signaling hyperactivation hallmark in colon cancer stem cells [171], and the regulating effect of this pathway in colorectal cancer, mediated by lncRNAs [181-183] made interesting to know whether a lncRNAs, in this case *LUST*, could act as a transcriptional

regulator of *Wnt*-signaling during colonospheres formation. Culture under FGF positive and EGF positive serum-free conditions has been used to expand stem-like spheroid cells from primary colorectal cancer cells (CRC) or from primary cell lines [20,215,216]. To address the functional role of *LUST* on cancer stem cell maintenance, I performed *in vitro* colonospheres forming assays and analyzed by qRT-PCR the lncRNA and stemness markers mRNA levels. The expression of CD24 and CD44 resulted increased when HT-29 cells are grown in suspension, as confirmation of the cell stemness potential and the enrichment of CCICs within the colonospheres [170]. Interestingly, the expression of *LUST* that was showed to be increased in my RNA-seq data obtained from colonospheres, is validated with this experiment. More directly, the increase of *LUST* is strongly significant after only 14 days in culture, when the stemness markers levels are not increased yet, and reaches a 10 fold change increase when cells are grown for 5 weeks in the same conditions (**Fig. 9.3**). This result suggests that *LUST* elevation could be an early event in the colonospheres formation process. Epithelial stem and progenitor cells can grow as spheroid aggregates and, upon growth factor removal and in presence of serum or extracellular matrix, can differentiate [20,195-197]. To test this ability, and to investigate *LUST* behavior in such conditions, I induced colonospheres differentiation by adding fetal bovine serum (FBS) and by removing growth factor from the culture medium [23]. As expected, when colonospheres differentiation is generated upon adding FBS, both mRNA and protein expression levels of the "stem-like" markers CD24 and CD44 is abrogated (**Fig. 10.3** and **Fig. 11.3**), confirming reduced stem-like potential. Interestingly, *LUST* expression strongly decreases upon induced differentiation, and this could be considered as a clear evidence of its involvement in stemness phenotype (**Fig. 10.3**). Recent studies have identified many lncRNA transcripts to be differentially expressed in pluripotent cells, both in ESCs and in iPSCs compared to fibroblasts or neuronal progenitors [101,217-219]. In order to investigate *LUST* involvement in *Wnt*-signaling regulation I performed loss of function experiments using

lock nucleic acids (LNA) antisense probes (**Fig. 12.3 A**), demonstrating that *LUST* inhibition impairs the transcriptional activation of *Wnt*-signaling target genes (**Fig. 13.3**). In particular, the reduction of *AXIN2, CCND1, MYC* and of the stemness markers *CD24* and *CD44* mRNA levels, suggest the loss of stemness character; moreover the reduction of *CCND1* and *MYC* gene product levels (**Fig. 14.3**) strongly indicate a reduction of cell proliferation (**Fig.12.3 B**) [87,88]. Of interest, *LUST* knockdown reduces active β-catenin, confirming the decrease of *Wnt*-signaling activation in absence of *LUST* (Fig. **13.3**). On the contrary, a dual luciferase assay was performed on *LUST* overexpressing cells, showing that the lncRNA alone as well as β-catenin is able to activate the TCF-4 reporter mini-gene. Interestingly, co-transfection of pcDNA3-*LUST* and pcDNA3-β-catenin induces increase in luciferase activity (**Fig. 15.3**). These results demonstrate that the *LUST* transcript and β-catenin coordinately regulate *Wnt*/β-catenin signaling. Furthermore, *LUST* overexpression in HT-29 colon cancer cell line, induces *Wnt*-signaling target genes and stemness markers transcription activation (**Fig. 16.3**), and the increase is also reflected at the protein level (**Fig. 17.3**), indicating that *LUST* actively regulates *Wnt*//β-catenin signaling. To identify the molecular basis by which the lncRNA regulates this pathway, I visualized *LUST* transcript subcellular localization to provide clues to its function. The nuclear/cytoplasmic RNA fractionation (**Fig. 18.3**) and the RNA fluorescence *in situ* hybridization (FISH) (**Fig. 19.3**) experiments revealed that *LUST* is a prominently nuclear lncRNA, to conceptualize a transcriptional or chromatin-based role. Previous studies reported that members of *Wnt* signaling such as β-catenin, can selectively bind RNA [220,221]. Edwards et al., showed that the *Armadillo* repeats found within β-catenin are an evolutionary conserved structural order as a helical repeat protein family, that closely corresponds with *Puf* repeats of the RNA binding protein Pumilio. The *Armadillo* repeat structure reveals an extended, rainbow shaped molecule, with tandem helical repeats, that bears unexpected resemblance to *Puf* repeats in Pumilio as a distinct structure with strong

affinity to 3'-UTRs in mRNA [220].

Figure 1.4: β-catenin and Pumilio protein structure

β-catenin is a member of the highly conserved helical repeat protein family through the Armadillo repeats. Two members of the family, β-catenin with *arm* repeats (*left*) is shown alongside Pumilio with *Puf* repeats (*right*). Shown below is a single repeat from each structure, aligned with functionally equivalent helices—H3 for *Arm* and *Puf* repeat structures. (*Adapted from* [220]).

This hypothesis, was further demonstrated by Kim et al. showing that β-catenin recognizes a specific RNA motif in the cyclooxygenase-2 mRNA 3'-UTR [221]. Based on these findings I sought to analyze the hypothetical interaction of *LUST* with *Wnt*/β-catenin pathway members, by immuno-precipitating β-catenin trough RNA-Immuno-Precipitation (RIP). The strong enrichment of *LUST* binding to this protein, compared to the negative control (**Fig. 20.3**), opened new hypothesis of whether this lncRNA could bind to a mutated isoforms of the same protein. The CaCo-2 colon cancer cell line harbors both, APC and β-catenin mutations [187], making this colorectal cancer cell line suitable for my purposes. In attempt of identifying new binding partners of *LUST*, in case the missense mutation could have abrogated the binding site of the lncRNA for β-catenin, I performed cross-linking immuno-precipitation of β-catenin and the transcription factor TCF-4, known to induce transcriptional activation of *Wnt*-target genes [59,60,222]. My results showed that: (i) *LUST* binds to β-catenin in CaCo-2 colorectal cancer cell

line (**Fig. 21.3**), therefore the missense mutation does not abrogate the binding site, instead, this mutation that lies within the third *Armadillo* repeat would increase the accessible surface area at that site [223], thus facilitating *LUST* binding to the protein; (ii) the increase of *LUST* binding is also reflected by RIP experiment performed in the same cell line (**Fig. 22.3**); (iii) lncRNA *LUST* binds to both, β-catenin and TCF-4, enforcing the hypothesis of a functional role in regulating *Wnt*-signaling (**Fig. 21.3**). Although these experiments add confirmations to my hypothesis, to further determine the regulatory role of *LUST* in *Wnt*/β-catenin signaling during colonospheres formation, the fallowing experiment was carried out, to examine whether overexpression of the lncRNA would eventually affect the process. The earlier colonospheres formation obtained (**Fig. 23.3**) confers the last and additional demonstration of the extremely important role of *LUST* in CICCs maintenance by regulating *Wnt*-signaling activation. The relevance of my thesis, therefore, results is in three distinct and significant areas of *LUST*-mediated processes in CCICs maintenance:

(i) My work represents the first detailed characterization of *LUST* localization and role in the maintenance of CCICs self-renewal.

(ii) I established a transcriptional activating regulatory model, whereby *LUST* function is essential for the transcriptional activation of the *Wnt*-signaling target genes (**Fig. 2.4**).

(iii) I show a novel *LUST* function in regulating *Wnt*-signaling, essential to insure the CCICs maintenance.

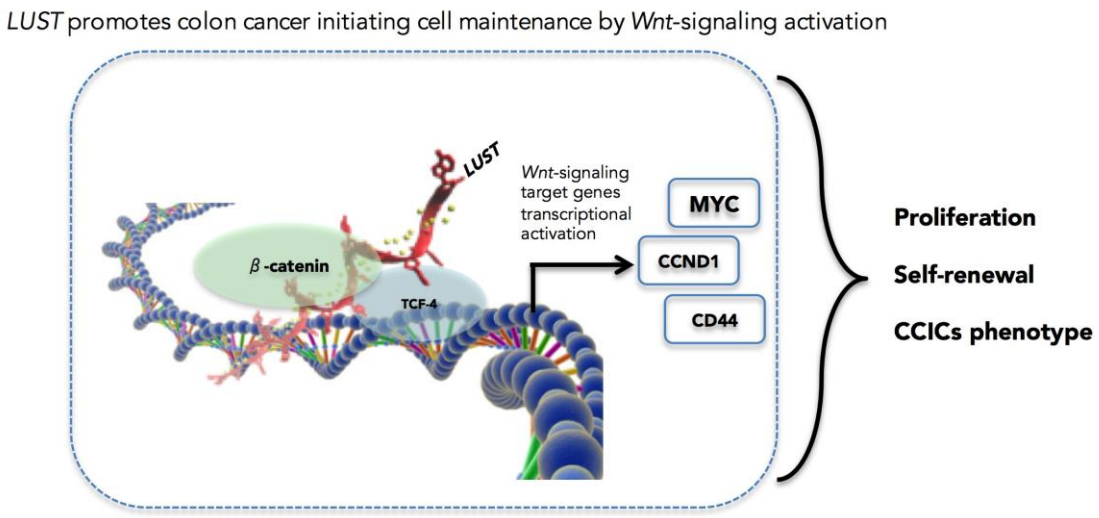

Figure 2.4: *LUST* acts as trascritpional regulator of *Wnt*-signaling

LUST regulates *Wnt*-signaling in CCICs through a coordinated physical interaction with β-catenin, inducing transcriptional activation of *Wnt*-signaling target genes, therefore promoting colon cancer initiating cells proliferation, self-renewal and phenotypic markers expression.

To conclude, I demonstrated that the lncRNA *LUST* is critical for CCICs self-renewal, by transcriptional regulating the expression of *Wnt*-signaling target genes. Overexpression of the *LUST* transcript increases the expression levels of *CCND1*, active β-catenin and other key regulators of cell cycle and promotes cell proliferation, thereby triggering colonospheres formation. The direct connection between *LUST* and β-catenin in maintaining CCICs self-renewal provides fundamental new insight into the transcriptional gene regulation network in cancer stem cell biology.

Given the importance of β-catenin in tumorigenesis, it is plausible to consider that *LUST* overexpression is involved with cancer cell fate. Therefore, dissecting the molecular mechanisms that mediate this overexpression envision a use in predicting cancer risk, achieve an early diagnosis, or track the prognosis of tumor fate, and thereby, provide a scope of diagnostic or therapeutic possibilities. Currently, nucleic acid-based methods prevail in targeting RNA, either by regulating the level of lncRNAs in cancer cells as well as modifying their structures or mature sequences. Among them, RNA interference (RNAi) based

techniques are arguably the most popular methods to inhibit lncRNAs in cancer cells. With the novel design of RNA through chemical modifications, the stability of these nucleic acid drugs is greatly improved [224], and combined with gene therapy technology, stable and persistent inhibition of aberrant lncRNAs can be resulted. Meanwhile, other established methods in inhibiting cancer-associated RNA, including antisense oligonucleotide (ASO), ribozyme and aptamer, are also effective to modulate lncRNAs, and they show unique features that can have advantages over siRNAs. On the other hand, the potential of the pharmacological modulation of β-catenin in cancer therapeutics may possibly provide an attractive option of targeting various aspects of the carcinogenic process i.e. initiation, progression and chemo-resistance in conjunction with the traditional chemotherapy. However, the long-term effects of the pharmacological manipulation of β-catenin remain still unclear and need to be taken in consideration, given that the overall regulation of this protein involves multiple signaling pathways. For this reason, the pharmacological modulation should be counterbalanced through the activation of compensatory signaling pathways. Furthermore, the possibility of adverse side effects of β-catenin inhibition cannot be ruled out at this juncture. To date, the use of various small molecule inhibitors of β-catenin targeting cancer have provided some encouraging results [225]. To this purpose, more efforts can be directed towards evaluating the efficacy of the existing inhibitors against colon cancer initiating cells and chemoresistant cancers, as it is evident that microenvironmental regulation of the β-catenin activity plays a central role in the malignant transformation and induction of metastasis; thus, in this case, the development of multiple approaches is key to identify the most effective lncRNA-based therapy. Therefore, *LUST* would represent a perfect candidate for developing lead compounds directed against colon cancer initiating cells. In fact, targeting *LUST*, hence the *Wnt*/β-catenin signaling activity, could open new avenues for novel and tailor-made cancer therapeutic approaches.

REFERENCES

1. Ricci-Vitiani, L., Fabrizi, E., Palio, E. & De Maria, R. Colon cancer stem cells. *J Mol Med (Berl)* **87**, 1097-104 (2009).

2. Boman, B.M. & Huang, E. Human colon cancer stem cells: a new paradigm in gastrointestinal oncology. *J Clin Oncol* **26**, 2828-38 (2008).

3. Medema, J.P. & Vermeulen, L. Microenvironmental regulation of stem cells in intestinal homeostasis and cancer. *Nature* **474**, 318-26 (2011).

4. Vries, R.G., Huch, M. & Clevers, H. Stem cells and cancer of the stomach and intestine. *Mol Oncol* **4**, 373-84 (2010).

5. Papailiou, J., Bramis, K.J., Gazouli, M. & Theodoropoulos, G. Stem cells in colon cancer. A new era in cancer theory begins. *Int J Colorectal Dis* **26**, 1-11 (2011).

6. Todaro, M., Francipane, M.G., Medema, J.P. & Stassi, G. Colon cancer stem cells: promise of targeted therapy. *Gastroenterology* **138**, 2151-62 (2010).

7. Dalerba, P., Cho, R.W. & Clarke, M.F. Cancer stem cells: models and concepts. *Annu Rev Med* **58**, 267-84 (2007).

8. Davies, E.J., Marsh, V. & Clarke, A.R. Origin and maintenance of the intestinal cancer stem cell. *Mol Carcinog* **50**, 254-63 (2011).

9. Potten, C.S., Gandara, R., Mahida, Y.R., Loeffler, M. & Wright, N.A. The stem cells of small intestinal crypts: where are they? *Cell Prolif* **42**, 731-50 (2009).

10. Vogelstein, B. et al. Genetic alterations during colorectal-tumor development. *N Engl J Med* **319**, 525-32 (1988).

11. Humphries, A. et al. Lineage tracing reveals multipotent stem cells maintain human adenomas and the pattern of clonal expansion in tumor evolution. *Proc Natl Acad Sci U S A* **110**, E2490-9 (2013).

12. Clevers, H. The cancer stem cell: premises, promises and challenges. *Nat Med* **17**, 313-9 (2011).

13. Furth, J.K., M. C. . The transmission of leukaemia of mice with a single cell. *Am J. Cancer* **31**, 276–282 (1937).

14. Belanger, L.F. & Leblond, C.P. A method for locating radioactive elements in tissues by covering histological sections with a photographic emulsion. *Endocrinology* **39**, 8-13 (1946).

15. Pierce, G.B. & Wallace, C. Differentiation of malignant to benign cells. *Cancer Res* **31**, 127-34 (1971).

16. Nowell, P.C. The clonal evolution of tumor cell populations. *Science* **194**, 23-8 (1976).

17. Lapidot, T. et al. A cell initiating human acute myeloid leukaemia after transplantation into SCID mice. *Nature* **367**, 645-8 (1994).

18. Bonnet, D. & Dick, J.E. Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. *Nat Med* **3**, 730-7 (1997).

19. Al-Hajj, M., Wicha, M.S., Benito-Hernandez, A., Morrison, S.J. & Clarke, M.F. Prospective identification of tumorigenic breast cancer cells. *Proc Natl Acad Sci U S A* **100**, 3983-8 (2003).

20. Singh, S.K. et al. Identification of human brain tumour initiating cells. *Nature* **432**, 396-401 (2004).

21. Dalerba, P. et al. Phenotypic characterization of human colorectal cancer stem cells. *Proc Natl Acad Sci U S A* **104**, 10158-63 (2007).

22. O'Brien, C.A., Pollett, A., Gallinger, S. & Dick, J.E. A human colon cancer cell capable of initiating tumour growth in immunodeficient mice. *Nature* **445**, 106-10 (2007).

23. Ricci-Vitiani, L. et al. Identification and expansion of human colon-cancer-initiating cells. *Nature* **445**, 111-5 (2007).

24. Dawood, S., Austin, L. & Cristofanilli, M. Cancer Stem Cells: Implications for Cancer Therapy. *Oncology (Williston Park)* **28**(2014).

25. O'Flaherty, J.D. et al. The cancer stem-cell hypothesis: its emerging role in lung cancer biology and its relevance for future therapy. *J Thorac Oncol* **7**, 1880-90 (2012).

26. Todaro, M. et al. Colon cancer stem cells dictate tumor growth and resist cell death by production of interleukin-4. *Cell Stem Cell* **1**, 389-402 (2007).

27. Vermeulen, L. et al. Single-cell cloning of colon cancer stem cells reveals a multi-lineage differentiation capacity. *Proc Natl Acad Sci U S A* **105**, 13427-32 (2008).

28. Karbanova, J. et al. The stem cell marker CD133 (Prominin-1) is expressed in various human glandular epithelia. *J Histochem Cytochem* **56**, 977-93 (2008).

29. Shmelkov, S.V. et al. CD133 expression is not restricted to stem cells, and both CD133+ and CD133- metastatic colon cancer cells initiate tumors. *J Clin Invest* **118**, 2111-20 (2008).

30. Du, L. et al. CD44 is of functional importance for colorectal cancer stem cells. *Clin Cancer Res* **14**, 6751-60 (2008).

31. Chu, P. et al. Characterization of a subpopulation of colon cancer cells with stem cell-like properties. *Int J Cancer* **124**, 1312-21 (2009).

32. Barker, N. et al. Crypt stem cells as the cells-of-origin of intestinal cancer. *Nature* **457**, 608-11 (2009).

33. Yeung, T.M., Gandhi, S.C., Wilding, J.L., Muschel, R. & Bodmer, W.F. Cancer stem cells from colorectal cancer-derived cell lines. *Proc Natl Acad Sci U S A* **107**, 3722-7 (2010).

34. Haraguchi, N. et al. CD133+CD44+ population efficiently enriches colon cancer initiating cells. *Ann Surg Oncol* **15**, 2927-33 (2008).

35. Pirruccello, S.J. & LeBien, T.W. The human B cell-associated antigen CD24 is a single chain sialoglycoprotein. *J Immunol* **136**, 3779-84 (1986).

36. Lee, H.J. et al. CD24, a novel cancer biomarker, predicting disease-free survival of non-small cell lung carcinomas: a retrospective study of prognostic factor analysis from the viewpoint of forthcoming (seventh) new TNM classification. *J Thorac Oncol* **5**, 649-57 (2010).

37. Kristiansen, G. et al. CD24 expression is a new prognostic marker in breast cancer. *Clin Cancer Res* **9**, 4906-13 (2003).

38. Pierres, M. et al. Evidence that murine hematopoietic cell subset marker J11d is attached to a glycosyl-phosphatidylinositol membrane anchor. *Eur J Immunol* **17**, 1781-5 (1987).

39. Springer, T., Galfre, G., Secher, D.S. & Milstein, C. Monoclonal xenogeneic antibodies to murine cell surface antigens: identification of novel leukocyte differentiation antigens. *Eur J Immunol* **8**, 539-51 (1978).

40. Fang, X., Zheng, P., Tang, J. & Liu, Y. CD24: from A to Z. *Cell Mol Immunol* **7**, 100-3 (2010).

41. Aigner, S. et al. CD24 mediates rolling of breast carcinoma cells on P-selectin. *FASEB J* **12**, 1241-51 (1998).

42. Zheng, J. et al. NDRG2 inhibits hepatocellular carcinoma adhesion, migration and invasion by regulating CD24 expression. *BMC Cancer* **11**, 251:1-9 (2011).

43. Taniuchi, K., Nishimori, I. & Hollingsworth, M.A. Intracellular CD24 inhibits cell invasion by posttranscriptional regulation of BART through interaction with G3BP. *Cancer Res* **71**, 895-905 (2011).

44. Visvader, J.E. & Lindeman, G.J. Cancer stem cells in solid tumours: accumulating evidence and unresolved questions. *Nat Rev Cancer* **8**, 755-68 (2008).

45. Sahlberg, S.H., Spiegelberg, D., Glimelius, B., Stenerlow, B. & Nestor, M. Evaluation of cancer stem cell markers CD133, CD44, CD24: association with AKT isoforms and radiation resistance in colon cancer cells. *PLoS One* **9**, e94621 (2014).

46. Sneath, R.J. & Mangham, D.C. The normal structure and function of CD44 and its role in neoplasia. *Mol Pathol* **51**, 191-200 (1998).

47. Ponta, H., Sherman, L. & Herrlich, P.A. CD44: from adhesion molecules to signalling regulators. *Nat Rev Mol Cell Biol* **4**, 33-45 (2003).

48. Naor, D., Wallach-Dayan, S.B., Zahalka, M.A. & Sionov, R.V. Involvement of CD44, a molecule with a thousand faces, in cancer dissemination. *Semin Cancer Biol* **18**, 260-7 (2008).

49. Rangaswami, H., Bulbule, A. & Kundu, G.C. Osteopontin: role in cell signaling and cancer progression. *Trends Cell Biol* **16**, 79-87 (2006).

50. Napier, S.L., Healy, Z.R., Schnaar, R.L. & Konstantopoulos, K. Selectin ligand expression regulates the initial vascular interactions of colon carcinoma cells: the roles

of CD44v and alternative sialofucosylated selectin ligands. *J Biol Chem* **282**, 3433-41 (2007).

51. Wielenga, V.J. et al. Expression of CD44 in Apc and Tcf mutant mice implies regulation by the WNT pathway. *Am J Pathol* **154**, 515-23 (1999).

52. Schmitt, M., Metzger, M., Gradl, D., Davidson, G. & Orian-Rousseau, V. CD44 functions in Wnt signaling by regulating LRP6 localization and activation. *Cell Death Differ* (2014).

53. Nusse, R. & Varmus, H.E. Many tumors induced by the mouse mammary tumor virus contain a provirus integrated in the same region of the host genome. *Cell* **31**, 99-109 (1982).

54. Rijsewijk, F. et al. The Drosophila homolog of the mouse mammary oncogene int-1 is identical to the segment polarity gene wingless. *Cell* **50**, 649-57 (1987).

55. Noordermeer, J., Klingensmith, J., Perrimon, N. & Nusse, R. dishevelled and armadillo act in the wingless signalling pathway in Drosophila. *Nature* **367**, 80-3 (1994).

56. Peifer, M., Sweeton, D., Casey, M. & Wieschaus, E. wingless signal and Zeste-white 3 kinase trigger opposing changes in the intracellular distribution of Armadillo. *Development* **120**, 369-80 (1994).

57. Siegfried, E., Chou, T.B. & Perrimon, N. wingless signaling acts through zeste-white 3, the Drosophila homolog of glycogen synthase kinase-3, to regulate engrailed and establish cell fate. *Cell* **71**, 1167-79 (1992).

58. McMahon, A.P. & Moon, R.T. Ectopic expression of the proto-oncogene int-1 in Xenopus embryos leads to duplication of the embryonic axis. *Cell* **58**, 1075-84 (1989).

59. Behrens, J. et al. Functional interaction of beta-catenin with the transcription factor LEF-1. *Nature* **382**, 638-42 (1996).

60. Molenaar, M. et al. XTcf-3 transcription factor mediates beta-catenin-induced axis formation in Xenopus embryos. *Cell* **86**, 391-9 (1996).

61. Kemler, R. From cadherins to catenins: cytoplasmic protein interactions and regulation of cell adhesion. *Trends Genet* **9**, 317-21 (1993).

62. Bhanot, P. et al. A new member of the frizzled family from Drosophila functions as a Wingless receptor. *Nature* **382**, 225-30 (1996).

63. Wehrli, M. et al. arrow encodes an LDL-receptor-related protein essential for Wingless signalling. *Nature* **407**, 527-30 (2000).

64. Kusserow, A. et al. Unexpected complexity of the Wnt gene family in a sea anemone. *Nature* **433**, 156-60 (2005).

65. Petersen, C.P. & Reddien, P.W. Wnt signaling and the polarity of the primary body axis. *Cell* **139**, 1056-68 (2009).

66. Tanaka, K., Kitagawa, Y. & Kadowaki, T. Drosophila segment polarity gene product porcupine stimulates the posttranslational N-glycosylation of wingless in the endoplasmic reticulum. *J Biol Chem* **277**, 12816-23 (2002).

67. Willert, K. et al. Wnt proteins are lipid-modified and can act as stem cell growth factors. *Nature* **423**, 448-52 (2003).

68. Franch-Marro, X., Wendler, F., Griffith, J., Maurice, M.M. & Vincent, J.P. In vivo role of lipid adducts on Wingless. *J Cell Sci* **121**, 1587-92 (2008).

69. Kurayoshi, M., Yamamoto, H., Izumi, S. & Kikuchi, A. Post-translational palmitoylation and glycosylation of Wnt-5a are necessary for its signalling. *Biochem J* **402**, 515-23 (2007).

70. Strand, M. & Micchelli, C.A. Quiescent gastric stem cells maintain the adult Drosophila stomach. *Proc Natl Acad Sci U S A* **108**, 17696-701 (2011).

71. Sato, T. et al. Paneth cells constitute the niche for Lgr5 stem cells in intestinal crypts. *Nature* **469**, 415-8 (2011).

72. Clevers, H. & Nusse, R. Wnt/beta-catenin signaling and disease. *Cell* **149**, 1192-205 (2012).

73. Dann, C.E. et al. Insights into Wnt binding and signalling from the structures of two Frizzled cysteine-rich domains. *Nature* **412**, 86-90 (2001).

74. Janda, C.Y., Waghray, D., Levin, A.M., Thomas, C. & Garcia, K.C. Structural basis of Wnt recognition by Frizzled. *Science* **337**, 59-64 (2012).

75. Pinson, K.I., Brennan, J., Monkley, S., Avery, B.J. & Skarnes, W.C. An LDL-receptor-related protein mediates Wnt signalling in mice. *Nature* **407**, 535-8 (2000).

76. Tamai, K. et al. LDL-receptor-related proteins in Wnt signal transduction. *Nature* **407**, 530-5 (2000).

77. Mao, J. et al. Low-density lipoprotein receptor-related protein-5 binds to Axin and regulates the canonical Wnt signaling pathway. *Mol Cell* **7**, 801-9 (2001).

78. Zeng, X. et al. A dual-kinase mechanism for Wnt co-receptor phosphorylation and activation. *Nature* **438**, 873-7 (2005).

79. Chen, W. et al. Dishevelled 2 recruits beta-arrestin 2 to mediate Wnt5A-stimulated endocytosis of Frizzled 4. *Science* **301**, 1391-4 (2003).

80. Fiedler, M., Mendoza-Topaz, C., Rutherford, T.J., Mieszczanek, J. & Bienz, M. Dishevelled interacts with the DIX domain polymerization interface of Axin to

interfere with its function in down-regulating beta-catenin. *Proc Natl Acad Sci U S A* **108**, 1937-42 (2011).

81. Schwarz-Romond, T. et al. The DIX domain of Dishevelled confers Wnt signaling by dynamic polymerization. *Nat Struct Mol Biol* **14**, 484-92 (2007).

82. Major, M.B. et al. Wilms tumor suppressor WTX negatively regulates WNT/beta-catenin signaling. *Science* **316**, 1043-6 (2007).

83. Aberle, H., Bauer, A., Stappert, J., Kispert, A. & Kemler, R. beta-catenin is a target for the ubiquitin-proteasome pathway. *Embo j* **16**, 3797-804 (1997).

84. Li, V.S. et al. Wnt signaling through inhibition of beta-catenin degradation in an intact Axin1 complex. *Cell* **149**, 1245-56 (2012).

85. Hikasa, H. et al. Regulation of TCF3 by Wnt-dependent phosphorylation during vertebrate axis specification. *Dev Cell* **19**, 521-32 (2010).

86. Lee, W., Swarup, S., Chen, J., Ishitani, T. & Verheyen, E.M. Homeodomain-interacting protein kinases (Hipks) promote Wnt/Wg signaling through stabilization of beta-catenin/Arm and stimulation of target gene expression. *Development* **136**, 241-51 (2009).

87. He, T.C. et al. Identification of c-MYC as a target of the APC pathway. *Science* **281**, 1509-12 (1998).

88. Tetsu, O. & McCormick, F. Beta-catenin regulates expression of cyclin D1 in colon carcinoma cells. *Nature* **398**, 422-6 (1999).

89. Goentoro, L. & Kirschner, M.W. Evidence that fold-change, and not absolute level, of beta-catenin dictates Wnt signaling. *Mol Cell* **36**, 872-84 (2009).

90. Maeda, N. et al. Transcript annotation in FANTOM3: mouse gene catalog based on physical cDNAs. *PLoS Genet* **2**, e62 (2006).

91. Djebali, S. et al. Landscape of transcription in human cells. *Nature* **489**, 101-8 (2012).

92. Amaral, P.P., Dinger, M.E., Mercer, T.R. & Mattick, J.S. The eukaryotic genome as an RNA machine. *Science* **319**, 1787-9 (2008).

93. Mattick, J.S. & Makunin, I.V. Non-coding RNA. *Hum Mol Genet* **15 Spec No 1**, R17-29 (2006).

94. Rinn, J.L. & Chang, H.Y. Genome regulation by long noncoding RNAs. *Annu Rev Biochem* **81**, 145-66 (2012).

95. Derrien, T. et al. The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. *Genome Res* **22**, 1775-89 (2012).

96. Batista, P.J. & Chang, H.Y. Long noncoding RNAs: cellular address codes in development and disease. *Cell* **152**, 1298-307 (2013).

97. Bernstein, B.E. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57-74 (2012).

98. Mercer, T.R., Dinger, M.E. & Mattick, J.S. Long non-coding RNAs: insights into functions. *Nat Rev Genet* **10**, 155-9 (2009).

99. Tsai, M.C., Spitale, R.C. & Chang, H.Y. Long intergenic noncoding RNAs: new links in cancer progression. *Cancer Res* **71**, 3-7 (2011).

100. Volders, P.J. et al. LNCipedia: a database for annotated human lncRNA transcript sequences and structures. *Nucleic Acids Res* **41**, D246-51 (2013).

101. Guttman, M. et al. Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature* **458**, 223-7 (2009).

102. Qureshi, I.A. & Mehler, M.F. Emerging roles of non-coding RNAs in brain evolution, development, plasticity and disease. *Nat Rev Neurosci* **13**, 528-41 (2012).

103. Huarte, M. et al. A large intergenic noncoding RNA induced by p53 mediates global gene repression in the p53 response. *Cell* **142**, 409-19 (2010).

104. Hiller, M. et al. Conserved introns reveal novel transcripts in Drosophila melanogaster. *Genome Res* **19**, 1289-300 (2009).

105. Wang, K.C. & Chang, H.Y. Molecular mechanisms of long noncoding RNAs. *Mol Cell* **43**, 904-14 (2011).

106. Yang, G., Lu, X. & Yuan, L. LncRNA: a link between RNA and cancer. *Biochim Biophys Acta* **1839**, 1097-109 (2014).

107. Pasic, I. et al. Recurrent focal copy-number changes and loss of heterozygosity implicate two noncoding RNAs and one tumor suppressor gene at chromosome 3q13.31 in osteosarcoma. *Cancer Res* **70**, 160-71 (2010).

108. Poliseno, L. et al. Deletion of PTENP1 pseudogene in human melanoma. *J Invest Dermatol* **131**, 2497-500 (2011).

109. Akrami, R. et al. Comprehensive analysis of long non-coding RNAs in ovarian cancer reveals global patterns and targeted DNA amplification. *PLoS One* **8**, e80306 (2013).

110. Chung, S. et al. Association of a novel long non-coding RNA in 8q24 with prostate cancer susceptibility. *Cancer Sci* **102**, 245-52 (2011).

111. Autuoro, J.M., Pirnie, S.P. & Carmichael, G.G. Long noncoding RNAs in imprinting and X chromosome inactivation. *Biomolecules* **4**, 76-100 (2014).

112. Chen, B. et al. Mdig de-represses H19 large intergenic non-coding RNA (lincRNA) by down-regulating H3K9me3 and heterochromatin. *Oncotarget* **4**, 1427-37 (2013).

113. Higashimoto, K., Soejima, H., Saito, T., Okumura, K. & Mukai, T. Imprinting disruption of the CDKN1C/KCNQ1OT1 domain: the molecular mechanisms causing Beckwith-Wiedemann syndrome and cancer. *Cytogenet Genome Res* **113**, 306-12 (2006).

114. Wijnen, M., Alders, M., Zwaan, C.M., Wagner, A. & van den Heuvel-Eibrink, M.M. KCNQ1OT1 hypomethylation: a novel disguised genetic predisposition in sporadic pediatric adrenocortical tumors? *Pediatr Blood Cancer* **59**, 565-6 (2012).

115. Weksberg, R. et al. Tumor development in the Beckwith-Wiedemann syndrome is associated with a variety of constitutional molecular 11p15 alterations including imprinting defects of KCNQ1OT1. *Hum Mol Genet* **10**, 2989-3000 (2001).

116. Nakano, S. et al. Expression profile of LIT1/KCNQ1OT1 and epigenetic status at the KvDMR1 in colorectal cancers. *Cancer Sci* **97**, 1147-54 (2006).

117. Hanahan, D. & Weinberg, R.A. Hallmarks of cancer: the next generation. *Cell* **144**, 646-74 (2011).

118. Tang, D.G. Understanding cancer stem cell heterogeneity and plasticity. *Cell Res* **22**, 457-72 (2012).

119. Trimarchi, T. et al. Genome-wide mapping and characterization of Notch-regulated long noncoding RNAs in acute leukemia. *Cell* **158**, 593-606 (2014).

120. Richly, H., Aloia, L. & Di Croce, L. Roles of the Polycomb group proteins in stem cells and cancer. *Cell Death Dis* **2**, e204 (2011).

121. Khalil, A.M. et al. Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. *Proc Natl Acad Sci U S A* **106**, 11667-72 (2009).

122. Yap, K.L. et al. Molecular interplay of the noncoding RNA ANRIL and methylated histone H3 lysine 27 by polycomb CBX7 in transcriptional silencing of INK4a. *Mol Cell* **38**, 662-74 (2010).

123. Aguilo, F., Zhou, M.M. & Walsh, M.J. Long noncoding RNA, polycomb, and the ghosts haunting INK4b-ARF-INK4a expression. *Cancer Res* **71**, 5365-9 (2011).

124. Gupta, R.A. et al. Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature* **464**, 1071-6 (2010).

125. Sheik Mohamed, J., Gaughwin, P.M., Lim, B., Robson, P. & Lipovich, L. Conserved long noncoding RNAs transcriptionally regulated by Oct4 and Nanog modulate pluripotency in mouse embryonic stem cells. *Rna* **16**, 324-37 (2010).

126. Chiou, S.H. et al. Coexpression of Oct4 and Nanog enhances malignancy in lung adenocarcinoma by inducing cancer stem cell-like properties and epithelial-mesenchymal transdifferentiation. *Cancer Res* **70**, 10433-44 (2010).

127. Prensner, J.R. & Chinnaiyan, A.M. The emergence of lncRNAs in cancer biology. *Cancer Discov* **1**, 391-407 (2011).

128. Fu, X., Ravindranath, L., Tran, N., Petrovics, G. & Srivastava, S. Regulation of apoptosis by a prostate-specific and prostate cancer-associated noncoding gene, PCGEM1. *DNA Cell Biol* **25**, 135-41 (2006).

129. Petrovics, G. et al. Elevated expression of PCGEM1, a prostate-specific gene with cell growth-promoting function, is associated with high-risk prostate cancer patients. *Oncogene* **23**, 605-11 (2004).

130. Srikantan, V. et al. PCGEM1, a prostate-specific gene, is overexpressed in prostate cancer. *Proc Natl Acad Sci U S A* **97**, 12216-21 (2000).

131. Matouk, I.J. et al. The H19 non-coding RNA is essential for human tumor growth. *PLoS One* **2**, e845 (2007).

132. Barsyte-Lovejoy, D. et al. The c-Myc oncogene directly induces the H19 noncoding RNA by allele-specific binding to potentiate tumorigenesis. *Cancer Res* **66**, 5330-7 (2006).

133. Mourtada-Maarabouni, M., Pickard, M.R., Hedge, V.L., Farzaneh, F. & Williams, G.T. GAS5, a non-protein-coding RNA, controls apoptosis and is downregulated in breast cancer. *Oncogene* **28**, 195-208 (2009).

134. Michalik, K.M. et al. Long noncoding RNA MALAT1 regulates endothelial cell function and vessel growth. *Circ Res* **114**, 1389-97 (2014).

135. Lai, M.C. et al. Long non-coding RNA MALAT-1 overexpression predicts tumor recurrence of hepatocellular carcinoma after liver transplantation. *Med Oncol* **29**, 1810-6 (2012).

136. Schmidt, L.H. et al. The long noncoding MALAT-1 RNA indicates a poor prognosis in non-small cell lung cancer and induces migration and tumor growth. *J Thorac Oncol* **6**, 1984-92 (2011).

137. Xu, C., Yang, M., Tian, J., Wang, X. & Li, Z. MALAT-1: a long non-coding RNA and its important 3' end functional motif in colorectal cancer metastasis. *Int J Oncol* **39**, 169-75 (2011).

138. Zhou, Y., Zhang, X. & Klibanski, A. MEG3 noncoding RNA: a tumor suppressor. *J Mol Endocrinol* **48**, R45-53 (2012).

139. Kim, K. et al. HOTAIR is a negative prognostic factor and exhibits pro-oncogenic activity in pancreatic cancer. *Oncogene* **32**, 1616-25 (2013).

140. Geng, Y.J., Xie, S.L., Li, Q., Ma, J. & Wang, G.Y. Large intervening non-coding RNA HOTAIR is associated with hepatocellular carcinoma progression. *J Int Med Res* **39**, 2119-28 (2011).

141. Niinuma, T. et al. Upregulation of miR-196a and HOTAIR drive malignant character in gastrointestinal stromal tumors. *Cancer Res* **72**, 1126-36 (2012).

142. Kogo, R. et al. Long noncoding RNA HOTAIR regulates polycomb-dependent chromatin modification and is associated with poor prognosis in colorectal cancers. *Cancer Res* **71**, 6320-6 (2011).

143. Yuan, J.H. et al. A long noncoding RNA activated by TGF-beta promotes the invasion-metastasis cascade in hepatocellular carcinoma. *Cancer Cell* **25**, 666-81 (2014).

144. Zhang, W. et al. A novel urinary long non-coding RNA transcript improves diagnostic accuracy in patients undergoing prostate biopsy. *Prostate* (2015).

145. Panzitt, K. et al. Characterization of HULC, a novel gene with striking up-regulation in hepatocellular carcinoma, as noncoding RNA. *Gastroenterology* **132**, 330-42 (2007).

146. Fradet, Y. et al. uPM3, a new molecular urine test for the detection of prostate cancer. *Urology* **64**, 311-5; discussion 315-6 (2004).

147. Shappell, S.B. Clinical utility of prostate carcinoma molecular diagnostic tests. *Rev Urol* **10**, 44-69 (2008).

148. Tong, Y.S. et al. Identification of the long non-coding RNA POU3F3 in plasma as a novel biomarker for diagnosis of esophageal squamous cell carcinoma. *Mol Cancer* **14**, 3 (2015).

149. Rintala-Maki, N.D. & Sutherland, L.C. Identification and characterisation of a novel antisense non-coding RNA from the RBM5 gene locus. *Gene* **445**, 7-16 (2009).

150. Sutherland, L.C. et al. LUCA-15-encoded sequence variants regulate CD95-mediated apoptosis. *Oncogene* **19**, 3774-81 (2000).

151. Shu, Y. et al. The apoptosis modulator and tumour suppressor protein RBM5 is a phosphoprotein. *Cell Biochem Funct* **25**, 643-53 (2007).

152. Sutherland, L.C., Rintala-Maki, N.D., White, R.D. & Morin, C.D. RNA binding motif (RBM) proteins: a novel family of apoptosis modulators? *J Cell Biochem* **94**, 5-24 (2005).

153. Rintala-Maki, N.D., Abrasonis, V., Burd, M. & Sutherland, L.C. Genetic instability of RBM5/LUCA-15/H37 in MCF-7 breast carcinoma sublines may affect susceptibility to apoptosis. *Cell Biochem Funct* **22**, 307-13 (2004).

154. Rintala-Maki, N.D. & Sutherland, L.C. LUCA-15/RBM5, a putative tumour suppressor, enhances multiple receptor-initiated death signals. *Apoptosis* **9**, 475-84 (2004).

155. Mourtada-Maarabouni, M., L. C. Sutherland, J. M. Meredith and G. T. Williams Simultaneous acceleration of the cell cycle and suppression of apoptosis by splice variant delta-6 of the candidate tumour suppressor LUCA-15/RBM5. *Genes Cells* **8**, 109-119 (2003).

156. Oh, J.J. et al. 3p21.3 tumor suppressor gene H37/Luca15/RBM5 inhibits growth of human lung cancer cells through cell cycle arrest and apoptosis. *Cancer Res* **66**, 3419-27 (2006).

157. Oh, J.J., West, A.R., Fishbein, M.C. & Slamon, D.J. A candidate tumor suppressor gene, H37, from the human lung cancer tumor suppressor locus 3p21.3. *Cancer Res* **62**, 3207-13 (2002).

158. Kreso, A. & O'Brien, C.A. Colon cancer stem cells. *Curr Protoc Stem Cell Biol* **Chapter 3**, Unit 3.1 (2008).

159. Langmead, B., Schatz, M.C., Lin, J., Pop, M. & Salzberg, S.L. Searching for SNPs with cloud computing. *Genome Biol* **10**, R134 (2009).

160. Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* **26**, 589-95 (2010).

161. Wang, L., Feng, Z., Wang, X., Wang, X. & Zhang, X. DEGseq: an R package for identifying differentially expressed genes from RNA-seq data. *Bioinformatics* **26**, 136-8 (2010).

162. Huang da, W., Sherman, B.T. & Lempicki, R.A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* **4**, 44-57 (2009).

163. Huang da, W. et al. Extracting biological meaning from large gene lists with DAVID. *Curr Protoc Bioinformatics* **Chapter 13**, Unit 13.11 (2009).

164. Barakat, T.S. & Gribnau, J. Combined DNA-RNA fluorescent in situ hybridization (FISH) to study X chromosome inactivation in differentiated female mouse embryonic stem cells. *J Vis Exp* (2014).

165. Moran, V.A., Niland, C.N. & Khalil, A.M. Co-Immunoprecipitation of long noncoding RNAs. *Methods Mol Biol* **925**, 219-28 (2012).

166. Kuo, M.H. & Allis, C.D. In vivo cross-linking and immunoprecipitation for studying dynamic Protein:DNA associations in a chromatin environment. *Methods* **19**, 425-33 (1999).

167. Ule, J., Jensen, K., Mele, A. & Darnell, R.B. CLIP: a method for identifying protein-RNA interaction sites in living cells. *Methods* **37**, 376-86 (2005).

168. Flatmark, K., Maelandsmo, G.M., Martinsen, M., Rasmussen, H. & Fodstad, O. Twelve colorectal cancer cell lines exhibit highly variable growth and metastatic capacities in an orthotopic model in nude mice. *Eur J Cancer* **40**, 1593-8 (2004).

169. van Klinken, B.J. et al. The human intestinal cell lines Caco-2 and LS174T as models to study cell-type specific mucin expression. *Glycoconj J* **13**, 757-68 (1996).

170. Kanwar, S.S., Yu, Y., Nautiyal, J., Patel, B.B. & Majumdar, A.P. The Wnt/beta-catenin pathway regulates growth and maintenance of colonospheres. *Mol Cancer* **9**, 212 (2010).

171. Vermeulen, L. et al. Wnt activity defines colon cancer stem cells and is regulated by the microenvironment. *Nat Cell Biol* **12**, 468-76 (2010).

172. Nilsson, H.E. et al. Intestinal MUC2 mucin supramolecular topology by packing and release resting on D3 domain assembly. *J Mol Biol* **426**, 2567-79 (2014).

173. Zhou, Q. et al. Keratin 20 helps maintain intermediate filament organization in intestinal epithelia. *Mol Biol Cell* **14**, 2959-71 (2003).

174. Huang, S. et al. Kruppel-like factor 9 inhibits glioma cell proliferation and tumorigenicity via downregulation of miR-21. *Cancer Lett* **356**, 547-55 (2015).

175. Palmer, H.G., Anjos-Afonso, F., Carmeliet, G., Takeda, H. & Watt, F.M. The vitamin D receptor is a Wnt effector that controls hair follicle differentiation and specifies tumor type in adult epidermis. *PLoS One* **3**, e1483 (2008).

176. Hu, L., Bikle, D.D. & Oda, Y. Reciprocal role of vitamin D receptor on beta-catenin regulated keratinocyte proliferation and differentiation. *J Steroid Biochem Mol Biol* **144 Pt A**, 237-41 (2014).

177. Jing, C. et al. Tazarotene-induced gene 1 (TIG1) expression in prostate carcinomas and its relationship to tumorigenicity. *J Natl Cancer Inst* **94**, 482-90 (2002).

178. Shukla, S. & Meeran, S.M. Epigenetics of cancer stem cells: Pathways and therapeutics. *Biochim Biophys Acta* **1840**, 3494-3502 (2014).

179. Chhabra, R. & Saini, N. MicroRNAs in cancer stem cells: current status and future directions. *Tumour Biol* **35**, 8395-405 (2014).

180. Eades, G. et al. Long non-coding RNAs in stem cells and cancer. *World J Clin Oncol* **5**, 134-41 (2014).

181. Wang, Q. et al. Exploring the Wnt Pathway-Associated LncRNAs and Genes Involved in Pancreatic Carcinogenesis Driven by Tp53 Mutation. *Pharm Res* **32**, 793-805 (2015).

182. Xiang, J.F. et al. Human colorectal cancer-specific CCAT1-L lncRNA regulates long-range chromatin interactions at the MYC locus. *Cell Res* **24**, 513-31 (2014).

183. Herriges, M.J. et al. Long noncoding RNAs are spatially correlated with transcription factors and regulate lung development. *Genes Dev* **28**, 1363-79 (2014).

184. Bernard, D. et al. A long nuclear-retained non-coding RNA regulates synaptogenesis by modulating gene expression. *Embo j* **29**, 3082-93 (2010).

185. Wang, G. et al. LincRNA-p21 enhances the sensitivity of radiotherapy for human colorectal cancer by targeting the Wnt/beta-catenin signaling pathway. *Oncol Rep* **31**, 1839-45 (2014).

186. Ji, Q. et al. Resveratrol inhibits invasion and metastasis of colorectal cancer cells via MALAT1 mediated Wnt/beta-catenin signal pathway. *PLoS One* **8**, e78700 (2013).

187. Ilyas, M., Tomlinson, I.P., Rowan, A., Pignatelli, M. & Bodmer, W.F. Beta-catenin mutations in cell lines established from human colorectal cancers. *Proc Natl Acad Sci U S A* **94**, 10330-4 (1997).

188. Korinek, V. et al. Two members of the Tcf family implicated in Wnt/beta-catenin signaling during embryogenesis in the mouse. *Mol Cell Biol* **18**, 1248-56 (1998).

189. Kaneko, S. et al. Interactions between JARID2 and noncoding RNAs regulate PRC2 recruitment to chromatin. *Mol Cell* **53**, 290-300 (2014).

190. Jung, P. et al. Isolation and in vitro expansion of human colonic stem cells. *Nat Med* **17**, 1225-7 (2011).

191. Huang, E.H. et al. Aldehyde dehydrogenase 1 is a marker for normal and malignant human colonic stem cells (SC) and tracks SC overpopulation during colon tumorigenesis. *Cancer Res* **69**, 3382-9 (2009).

192. Kemper, K. et al. Monoclonal antibodies against Lgr5 identify human colorectal cancer stem cells. *Stem Cells* **30**, 2378-86 (2012).

193. Todaro, M. et al. CD44v6 is a marker of constitutive and reprogrammed cancer stem cells driving colon cancer metastasis. *Cell Stem Cell* **14**, 342-56 (2014).

194. Kobayashi, S. et al. LGR5-positive colon cancer stem cells interconvert with drug-resistant LGR5-negative cells and are capable of tumor reconstitution. *Stem Cells* **30**, 2631-44 (2012).

195. Vescovi, A.L. et al. Isolation and cloning of multipotential stem cells from the embryonic human CNS and establishment of transplantable human neural stem cell lines by epigenetic stimulation. *Exp Neurol* **156**, 71-83 (1999).

196. Dontu, G. et al. In vitro propagation and transcriptional profiling of human mammary stem/progenitor cells. *Genes Dev* **17**, 1253-70 (2003).

197. Singh, S.K. et al. Identification of a cancer stem cell in human brain tumors. *Cancer Res* **63**, 5821-8 (2003).

198. Jin, L., Hope, K.J., Zhai, Q., Smadja-Joffe, F. & Dick, J.E. Targeting of CD44 eradicates human acute myeloid leukemic stem cells. *Nat Med* **12**, 1167-74 (2006).

199. Kryczek, I. et al. IL-22(+)CD4(+) T cells promote colorectal cancer stemness via STAT3 transcription factor activation and induction of the methyltransferase DOT1L. *Immunity* **40**, 772-84 (2014).

200. Lotti, F. et al. Chemotherapy activates cancer-associated fibroblasts to maintain colorectal cancer-initiating cells by IL-17A. *J Exp Med* **210**, 2851-72 (2013).

201. Kreso, A. & Dick, J.E. Evolution of the cancer stem cell model. *Cell Stem Cell* **14**, 275-91 (2014).

202. Langbein, L., Eckhart, L., Rogers, M.A., Praetzel-Wunder, S. & Schweizer, J. Against the rules: human keratin K80: two functional alternative splice variants, K80 and K80.1, with special cellular localization in a wide range of epithelia. *J Biol Chem* **285**, 36909-21 (2010).

203. O'Brien, C.A. et al. ID1 and ID3 regulate the self-renewal capacity of human colon cancer-initiating cells through p21. *Cancer Cell* **21**, 777-92 (2012).

204. Kreso, A. et al. Self-renewal as a therapeutic target in human colorectal cancer. *Nat Med* **20**, 29-36 (2014).

205. Zeuner, A., Todaro, M., Stassi, G. & De Maria, R. Colorectal cancer stem cells: from the crypt to the clinic. *Cell Stem Cell* **15**, 692-705 (2014).

206. Padua Alves, C. et al. Brief report: The lincRNA Hotair is required for epithelial-to-mesenchymal transition and stemness maintenance of cancer cell lines. *Stem Cells* **31**, 2827-32 (2013).

207. Ghosal, S., Das, S. & Chakrabarti, J. Long noncoding RNAs: new players in the molecular mechanism for maintenance and differentiation of pluripotent stem cells. *Stem Cells Dev* **22**, 2240-53 (2013).

208. Gutschner, T. & Diederichs, S. The hallmarks of cancer: a long non-coding RNA point of view. *RNA Biol* **9**, 703-19 (2012).

209. Nissan, A. et al. Colon cancer associated transcript-1: a novel RNA expressed in malignant and pre-malignant human tissues. *Int J Cancer* **130**, 1598-606 (2012).

210. Brocardo, M., Nathke, I.S. & Henderson, B.R. Redefining the subcellular location and transport of APC: new insights using a panel of antibodies. *EMBO Rep* **6**, 184-90 (2005).

211. Hinoi, T. et al. Complex formation of adenomatous polyposis coli gene product and axin facilitates glycogen synthase kinase-3 beta-dependent phosphorylation of beta-catenin and down-regulates beta-catenin. *J Biol Chem* **275**, 34399-406 (2000).

212. Nishisho, I. et al. Mutations of chromosome 5q21 genes in FAP and colorectal cancer patients. *Science* **253**, 665-9 (1991).

213. Scholer-Dahirel, A. et al. Maintenance of adenomatous polyposis coli (APC)-mutant colorectal cancer is dependent on Wnt/beta-catenin signaling. *Proc Natl Acad Sci U S A* **108**, 17135-40 (2011).

214. Tanaka, K. et al. Loss of imprinting of long QT intronic transcript 1 in colorectal cancer. *Oncology* **60**, 268-73 (2001).

215. Ieta, K. et al. Biological and genetic characteristics of tumor-initiating cells in colon cancer. *Ann Surg Oncol* **15**, 638-48 (2008).

216. Lombardo, Y. et al. Bone morphogenetic protein 4 induces differentiation of colorectal cancer stem cells and increases their response to chemotherapy in mice. *Gastroenterology* **140**, 297-309 (2011).

217. Loewer, S. et al. Large intergenic non-coding RNA-RoR modulates reprogramming of human induced pluripotent stem cells. *Nat Genet* **42**, 1113-7 (2010).

218. Ng, S.Y., Johnson, R. & Stanton, L.W. Human long non-coding RNAs promote pluripotency and neuronal differentiation by association with chromatin modifiers and transcription factors. *EMBO J* **31**, 522-33 (2012).

219. Kim, D.H. et al. Single-cell transcriptome analysis reveals dynamic changes in lncRNA expression during reprogramming. *Cell Stem Cell* **16**, 88-101 (2015).

220. Edwards, T.A., Pyle, S.E., Wharton, R.P. & Aggarwal, A.K. Structure of Pumilio reveals similarity between RNA and peptide binding motifs. *Cell* **105**, 281-9 (2001).

221. Kim, I., Kwak, H., Lee, H.K., Hyun, S. & Jeong, S. beta-Catenin recognizes a specific RNA motif in the cyclooxygenase-2 mRNA 3'-UTR and interacts with HuR in colon cancer cells. *Nucleic Acids Res* **40**, 6863-72 (2012).

222. Huber, O. et al. Nuclear localization of beta-catenin by interaction with transcription factor LEF-1. *Mech Dev* **59**, 3-10 (1996).

223. Ausubel, F.M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K., Eds. *Current Protocols in Molecular Biology* **1**, 4.7.1-4.7.6 (1993).

224. Geary, R.S. et al. Pharmacokinetic properties of 2'-O-(2-methoxyethyl)-modified oligonucleotide analogs in rats. *J Pharmacol Exp Ther* **296**, 890-7 (2001).

225. Thakur, R. & Mishra, D.P. Pharmacological modulation of beta-catenin and its applications in cancer therapy. *J Cell Mol Med* **17**, 449-56 (2013).

APPENDIX

Publications:

1. **F. Aguilo' *, S. Di Cecilia *** and M.J. Walsh. Long non-coding RNAs in Human Disease. Long non–coding RNA *ANRIL* and *Polycomb* in Human Cancers and Cardiovascular Disease. Book Chapter. Current Topics in Microbiology and Immunology; Springer. In Press. *Co-authors

2. F. Aguilo, F. Zhang*, A. Sancho*, **S. Di Cecilia**, A. Vashisht, Chih-Hung Chen, D.F. Lee, F. Jahouh, B. Andino, A. Roman, S. R. Krig, R. Wang, W. Zhang, J.A. Wohlschlegel & M. J. Walsh. Zinc Finger Protein 217 Regulates N^6-methyladenosine Deposition in Embryonic Stem Cell Transcripts. Under review for Cell Stem Cell.

3. **S. Di Cecilia**, A. Sancho, F. Zhang, F. Aguilo', M. Rengasamy, L. Del Vecchio, F. Salvatore, M.J. Walsh. The lncRNA *LUST* promotes Colon Cancer Initiating Cells maintenance by *Wnt*-signaling activation. Manuscript in preparation.

Proceedings:

1. **S. Di Cecilia**, F. Zhang, W. Zhang, L. Del Vecchio, F. Salvatore and M.J. Walsh. Genome-wide characterization of lncRNAs in human colon adenocarcinoma and colon cancer initiating cells, coupled with *Wnt* Signaling through β-catenin binding. Keystone Symposia. Santa Fe (NM),USA. Feb.27th- March 4th 2014

2. F. Aguilo', F. Zang, A.Vashisht, **S. Di Cecilia**, A. Sancho, C. Chung-hung, D.F. Lee, B. Andino, J. Wohlschlegel, M.J. Walsh. Comprehensive role of Zfp217 in m6A methylation. Cell Symposia: Regulatory RNAs, Berkeley, CA, USA. Oct 19-21 2014.

3. Scalia G., **Di Cecilia S.**, Raia M., Pascariello C., Abate G., Villani G., Visconte F., Mattioli M. and Del Vecchio L. Aberrant antigen expression (AEE) as a tool to identify patients with clonal myelodysplasia. XXIX National Conference of Cytometry. October 5th-8th 2011, Salerno (IT).

Courses:

May 2011

Biotecnologie e Medicina. Naples (IT).

Impact of the new Polychromatic Cytometry in Hematology. Naples (IT).

April 2011

RICS- Rete Italiana Cell Sorting-.Milan (IT).

March 2011

Teorical and pratical course: "**La citometria dell'emoglobinuria parossistica notturna (EPN)**. Naples (IT).